

Result No.	Score	Query	Match	Length	DB ID	Description
1	2454	AAG52573	21	466	100.0	Arabidopsis thaliana
2	2446	AAG52573	21	466	99.7	Arabidopsis thaliana
3	2446	AAG52573	21	485	99.7	Arabidopsis thaliana
4	2009	AAG52574	21	380	81.9	Arabidopsis thaliana
5	2001	AAG52574	21	380	81.5	Arabidopsis thaliana
6	1554	AAG52575	21	295	63.3	Arabidopsis thaliana
7	1246	AAG52575	21	440	50.8	Asparaginylendopeptidase
8	1246	AAG52575	21	490	50.8	Arabidopsis thaliana
9	1246	AAG52575	21	494	50.8	Asparaginylendopeptidase
10	1245	AAG52575	14	440	50.7	Asparaginylendopeptidase

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result ID	Score	Query ID	Match	Length	DB ID	Description
PD	06-SEP-2000	EP1033405-A2	XX	XX	XX	Arabidopsis thaliana
PF	25-FEB-2000	20000EP-0301439	XX	XX	PR	Arabidopsis thaliana
PR	25-FEB-1999	990US-0121825	XX	XX	PR	Arabidopsis thaliana
PR	05-MAR-1999	990US-0123180	XX	XX	PR	Arabidopsis thaliana
PR	09-MAR-1999	990US-0123548	XX	XX	PR	Arabidopsis thaliana
PR	23-MAR-1999	990US-0125788	XX	XX	PR	Arabidopsis thaliana
PR	29-MAR-1999	990US-0126264	XX	XX	PR	Arabidopsis thaliana
PR	01-APR-1999	990US-012785	XX	XX	PR	Arabidopsis thaliana
PR	06-APR-1999	990US-0128234	XX	XX	PR	Arabidopsis thaliana
PR	08-APR-1999	990US-0128714	XX	XX	PR	Arabidopsis thaliana
PR	16-APR-1999	990US-0129845	XX	XX	PR	Arabidopsis thaliana
PR	19-APR-1999	990US-0130077	XX	XX	PR	Arabidopsis thaliana

102(9)

ALIGNMENTS

RESULT 1
AAG52573
ID AAG52573 standard; Protein: 466 AA.
AC AAG52573;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 66847.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
PN EP1033405-A2.

PR	21-APR-1999;	990US-0130449;	PR	20-JUL-1999;	990US-0144352;
PR	23-APR-1999;	990US-0130510;	PR	20-JUL-1999;	990US-0144632;
PR	23-APR-1999;	990US-0130891;	PR	20-JUL-1999;	990US-0144884;
PR	28-APR-1999;	990US-0131449;	PR	21-JUL-1999;	990US-0144814;
PR	30-APR-1999;	990US-0132048;	PR	21-JUL-1999;	990US-0145086;
PR	04-MAY-1999;	990US-0132407;	PR	21-JUL-1999;	990US-0145088;
PR	05-MAY-1999;	990US-0132484;	PR	22-JUL-1999;	990US-0145087;
PR	06-MAY-1999;	990US-0132485;	PR	22-JUL-1999;	990US-0145089;
PR	06-MAY-1999;	990US-0132486;	PR	22-JUL-1999;	990US-0145192;
PR	07-MAY-1999;	990US-0132863;	PR	23-JUL-1999;	990US-0145145;
PR	11-MAY-1999;	990US-0134256;	PR	23-JUL-1999;	990US-0145218;
PR	14-MAY-1999;	990US-0134218;	PR	23-JUL-1999;	990US-0145224;
PR	14-MAY-1999;	990US-0134219;	PR	26-JUL-1999;	990US-0145276;
PR	14-MAY-1999;	990US-0134221;	PR	27-JUL-1999;	990US-0145913;
PR	14-MAY-1999;	990US-0134370;	PR	27-JUL-1999;	990US-0145918;
PR	18-MAY-1999;	990US-0134768;	PR	27-JUL-1999;	990US-0145919;
PR	19-MAY-1999;	990US-0134941;	PR	28-JUL-1999;	990US-0145951;
PR	20-MAY-1999;	990US-0135124;	PR	02-AUG-1999;	990US-0146386;
PR	21-MAY-1999;	990US-0135353;	PR	02-AUG-1999;	990US-0146388;
PR	24-MAY-1999;	990US-0135629;	PR	02-AUG-1999;	990US-0146389;
PR	25-MAY-1999;	990US-0136021;	PR	03-AUG-1999;	990US-0147038;
PR	27-MAY-1999;	990US-0136392;	PR	04-AUG-1999;	990US-014704;
PR	28-MAY-1999;	990US-0136782;	PR	04-AUG-1999;	990US-0147302;
PR	01-JUN-1999;	990US-0137222;	PR	05-AUG-1999;	990US-0147192;
PR	03-JUN-1999;	990US-0137528;	PR	05-AUG-1999;	990US-0147260;
PR	04-JUN-1999;	990US-0137502;	PR	06-AUG-1999;	990US-0147303;
PR	07-JUN-1999;	990US-0137724;	PR	06-AUG-1999;	990US-0147416;
PR	08-JUN-1999;	990US-0138094;	PR	09-AUG-1999;	990US-0147493;
PR	10-JUN-1999;	990US-0138447;	PR	09-AUG-1999;	990US-0147935;
PR	14-JUN-1999;	990US-0139119;	PR	10-AUG-1999;	990US-0148171;
PR	16-JUN-1999;	990US-0139452;	PR	11-AUG-1999;	990US-0148319;
PR	17-JUN-1999;	990US-0139453;	PR	12-AUG-1999;	990US-0148341;
PR	17-JUN-1999;	990US-0139492;	PR	13-AUG-1999;	990US-0148565;
PR	18-JUN-1999;	990US-0139454;	PR	13-AUG-1999;	990US-0148684;
PR	18-JUN-1999;	990US-0139455;	PR	16-AUG-1999;	990US-0149368;
PR	18-JUN-1999;	990US-0139456;	PR	17-AUG-1999;	990US-0149175;
PR	18-JUN-1999;	990US-0139457;	PR	18-AUG-1999;	990US-0149426;
PR	18-JUN-1999;	990US-0139458;	PR	20-AUG-1999;	990US-0149722;
PR	18-JUN-1999;	990US-0139459;	PR	20-AUG-1999;	990US-0149723;
PR	22-JUN-1999;	990US-0139459;	PR	20-AUG-1999;	990US-0149902;
PR	22-JUN-1999;	990US-0139460;	PR	23-AUG-1999;	990US-0149902;
PR	18-JUN-1999;	990US-0139461;	PR	16-AUG-1999;	990US-0149930;
PR	23-JUN-1999;	990US-0140353;	PR	25-AUG-1999;	990US-0150566;
PR	24-JUN-1999;	990US-0140354;	PR	26-AUG-1999;	990US-0150894;
PR	28-JUN-1999;	990US-0140695;	PR	27-AUG-1999;	990US-0151055;
PR	29-JUN-1999;	990US-0140823;	PR	13-SEP-1999;	990US-0153758;
PR	09-JUL-1999;	990US-0140991;	PR	13-SEP-1999;	990US-0154018;
PR	12-JUL-1999;	990US-0141287;	PR	15-SEP-1999;	990US-0154039;
PR	01-JUL-1999;	990US-0141842;	PR	16-SEP-1999;	990US-0154339;
PR	13-JUL-1999;	990US-0142154;	PR	20-SEP-1999;	990US-0154779;
PR	14-JUL-1999;	990US-0143624;	PR	05-OCT-1999;	990US-0157753;
PR	06-JUL-1999;	990US-0144055;	PR	06-OCT-1999;	990US-0157865;
PR	16-JUL-1999;	990US-0144085;	PR	07-OCT-1999;	990US-0158029;
PR	19-JUL-1999;	990US-0144086;	PR	08-OCT-1999;	990US-0158232;
PR	19-JUL-1999;	990US-0144325;	PR	12-OCT-1999;	990US-0159359;
PR	19-JUL-1999;	990US-0144331;	PR	13-OCT-1999;	990US-0159293;
PR	19-JUL-1999;	990US-0144333;	PR	13-OCT-1999;	990US-0159295;
PR	19-JUL-1999;	990US-0144334;	PR	14-OCT-1999;	990US-0159329;
PR	19-JUL-1999;	990US-0144335;	PR	14-OCT-1999;	990US-0159330.

PR	28-JUN-1999;	99US-0140823;
PR	29-JUN-1999;	99US-0140991;
PR	30-JUN-1999;	99US-0141287;
PR	01-JUL-1999;	99US-0141842;
PR	02-JUL-1999;	99US-0142154;
PR	02-JUL-1999;	99US-0142055;
PR	06-JUL-1999;	99US-0142390;
PR	08-JUL-1999;	99US-0142803;
PR	09-JUL-1999;	99US-0142920;
PR	12-JUL-1999;	99US-0142971;
PR	13-JUL-1999;	99US-0143342;
PR	14-JUL-1999;	99US-0143624;
PR	15-JUL-1999;	99US-0144005;
PR	16-JUL-1999;	99US-0144085;
PR	16-JUL-1999;	99US-0144086;
PR	19-JUL-1999;	99US-0144325;
PR	19-JUL-1999;	99US-0144331;
PR	19-JUL-1999;	99US-0144332;
PR	19-JUL-1999;	99US-0144333;
PR	19-JUL-1999;	99US-0144334;
PR	19-JUL-1999;	99US-0144335;
PR	20-JUL-1999;	99US-0144352;
PR	20-JUL-1999;	99US-0144632;
PR	20-JUL-1999;	99US-0144884;
PR	21-JUL-1999;	99US-0144814;
PR	21-JUL-1999;	99US-0145086;
PR	21-JUL-1999;	99US-0145088;
PR	22-JUL-1999;	99US-0145085;
PR	22-JUL-1999;	99US-0145087;
PR	22-JUL-1999;	99US-0145089;
PR	22-JUL-1999;	99US-0145192;
PR	23-JUL-1999;	99US-0145145;
PR	23-JUL-1999;	99US-0145218;
PR	23-JUL-1999;	99US-0145224;
PR	26-JUL-1999;	99US-0145216;
PR	27-JUL-1999;	99US-0145913;
PR	27-JUL-1999;	99US-0145918;
PR	27-JUL-1999;	99US-0145919;
PR	28-JUL-1999;	99US-0145951;
PR	02-AUG-1999;	99US-0146386;
PR	02-AUG-1999;	99US-0146388;
PR	03-AUG-1999;	99US-0145389;
PR	04-AUG-1999;	99US-0147204;
PR	04-AUG-1999;	99US-0147302;
PR	05-AUG-1999;	99US-0147192;
PR	05-AUG-1999;	99US-0147260;
PR	06-AUG-1999;	99US-0147303;
PR	06-AUG-1999;	99US-0147416;
PR	09-AUG-1999;	99US-0147493;
PR	09-AUG-1999;	99US-0147935;
PR	10-AUG-1999;	99US-0148171;
PR	11-AUG-1999;	99US-0148319;
PR	12-AUG-1999;	99US-0148411;
PR	13-AUG-1999;	99US-0148411;
PR	13-AUG-1999;	99US-0148565;
PR	16-AUG-1999;	99US-0148684;
PR	16-AUG-1999;	99US-0149368;
PR	17-AUG-1999;	99US-0149375;
PR	18-AUG-1999;	99US-0149226;
PR	20-AUG-1999;	99US-0149722;
PR	20-AUG-1999;	99US-0149723;
PR	20-AUG-1999;	99US-0149729;
PR	23-AUG-1999;	99US-0149902;
PR	23-AUG-1999;	99US-0149930;
PR	26-AUG-1999;	99US-0150566;
PR	27-AUG-1999;	99US-0151065;
PR	27-AUG-1999;	99US-0151066;
PR	30-AUG-1999;	99US-0151080;
PR	31-AUG-1999;	99US-0151303;
PR	01-SEP-1999;	99US-0151338;
PR	01-SEP-1999;	99US-0151930;
PR	01-SEP-1999;	99US-0152033;
PR	13-SEP-1999;	99US-0153788;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154019;
PR	20-SEP-1999;	99US-0155119;
PR	22-SEP-1999;	99US-0155117;
PR	23-SEP-1999;	99US-0155659;
PR	24-SEP-1999;	99US-0155785;
PR	28-SEP-1999;	99US-0156438;
PR	04-OCT-1999;	99US-0156596;
PR	05-OCT-1999;	99US-0157733;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158232;
PR	08-OCT-1999;	99US-0158369;
PR	12-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159295;
PR	14-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0161638;
PR	18-OCT-1999;	99US-0160841;
PR	21-OCT-1999;	99US-0160980;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160771;
PR	21-OCT-1999;	99US-0160814;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160982;
PR	22-OCT-1999;	99US-0160983;
PR	22-OCT-1999;	99US-0160984;
PR	22-OCT-1999;	99US-0160985;
PR	22-OCT-1999;	99US-0160986;
PR	22-OCT-1999;	99US-0160987;
PR	22-OCT-1999;	99US-0160988;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142.

QY 361 LESKEAQKKLDEKKNHRKQIDQSITDILRLSVKQTNVYLNLITSTRTRGQPLVDDWDCFKT 420
 Db 361 LESKEAQKKLDEKKNHRKQIDQSITDILRLSVKQTNVYLNLITSTRTRGQPLVDDWDCFKT 420

QY 421 LVNSFKNHCGATVHYGLKYTGALANICNMGVDKYQTVSAIEQACSM 466
 QY 421 LVNSFKNHCGATVHYGLKYTGALANICNMGVDKYQTVSAIEQACSM 466

Db 421 LVNSFKNHCGATVHYGLKYTGALANICNMGVDKYQTVSAIEQACSM 466

RESULT 3
 AAG12660 standard; Protein: 485 AA.
 ID AAG12660
 XX AC AAG12660;
 XX DT 17-CCT-2000 (first entry)
 XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 11860.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 hybridization assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX OS Arabidopsis thaliana.
 XX PN EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 25-FEB-1999; 990US-0121825.
 PR 05-MAR-1999; 990US-0123180.
 PR 09-MAR-1999; 990US-0123548.
 PR 23-MAR-1999; 990US-0125788.
 PR 25-MAR-1999; 990US-0126784.
 PR 29-MAR-1999; 990US-0126785.
 PR 01-APR-1999; 990US-0127462.
 PR 06-APR-1999; 990US-0128434.
 PR 08-APR-1999; 990US-0128714.
 PR 16-APR-1999; 990US-0129845.
 PR 19-APR-1999; 990US-0130077.
 PR 21-APR-1999; 990US-0130449.
 PR 23-APR-1999; 990US-0130510.
 PR 23-APR-1999; 990US-0130891.
 PR 28-APR-1999; 990US-0131449.
 PR 30-APR-1999; 990US-0132048.
 PR 30-APR-1999; 990US-0132487.
 PR 04-MAY-1999; 990US-0133484.
 PR 05-MAY-1999; 990US-0132485.
 PR 06-MAY-1999; 990US-0132486.
 PR 06-MAY-1999; 990US-0132487.
 PR 07-MAY-1999; 990US-0132863.
 PR 11-MAY-1999; 990US-0134256.
 PR 14-MAY-1999; 990US-0134218.
 PR 14-MAY-1999; 990US-0134219.
 PR 14-MAY-1999; 990US-0134221.
 PR 14-MAY-1999; 990US-0134370.
 PR 18-MAY-1999; 990US-0134768.
 PR 19-MAY-1999; 990US-0134911.
 PR 20-MAY-1999; 990US-0135124.
 PR 21-MAY-1999; 990US-0135353.
 PR 24-MAY-1999; 990US-0135629.
 PR 25-MAY-1999; 990US-0136021.
 PR 27-MAY-1999; 990US-0136392.
 PR 28-MAY-1999; 990US-0136782.
 PR 01-JUN-1999; 990US-0137222.
 PR 03-JUN-1999; 990US-0137528.
 PR 04-JUN-1999; 990US-0137502.
 PR 07-JUN-1999; 990US-0137724.
 PR 08-JUN-1999; 990US-0138094.
 PR 10-JUN-1999; 990US-0138540.
 PR 10-JUN-1999; 990US-0138847.

PR 14-JUN-1999; 990US-0139119.
 PR 16-JUN-1999; 990US-0139452.
 PR 16-JUN-1999; 990US-0139453.
 PR 17-JUN-1999; 990US-0139492.
 PR 18-JUN-1999; 990US-0139454.
 PR 18-JUN-1999; 990US-0139455.
 PR 18-JUN-1999; 990US-0139456.
 PR 18-JUN-1999; 990US-0139457.
 PR 18-JUN-1999; 990US-0139458.
 PR 18-JUN-1999; 990US-0139459.
 PR 18-JUN-1999; 990US-0139460.
 PR 18-JUN-1999; 990US-0139461.
 PR 18-JUN-1999; 990US-0139462.
 PR 18-JUN-1999; 990US-0139463.
 PR 18-JUN-1999; 990US-0139750.
 PR 18-JUN-1999; 990US-0139763.
 PR 21-JUN-1999; 990US-0139817.
 PR 22-JUN-1999; 990US-0139899.
 PR 23-JUN-1999; 990US-0140353.
 PR 24-JUN-1999; 990US-0140695.
 PR 28-JUN-1999; 990US-0140823.
 PR 29-JUN-1999; 990US-0140991.
 PR 30-JUN-1999; 990US-0141287.
 PR 01-JUL-1999; 990US-0141842.
 PR 01-JUL-1999; 990US-0142154.
 PR 02-JUL-1999; 990US-0142055.
 PR 06-JUL-1999; 990US-0142390.
 PR 08-JUL-1999; 990US-0142803.
 PR 09-JUL-1999; 990US-0142920.
 PR 12-JUL-1999; 990US-0142977.
 PR 13-JUL-1999; 990US-0143542.
 PR 14-JUL-1999; 990US-0143624.
 PR 15-JUL-1999; 990US-0144005.
 PR 16-JUL-1999; 990US-0144085.
 PR 16-JUL-1999; 990US-0144086.
 PR 19-JUL-1999; 990US-0144325.
 PR 19-JUL-1999; 990US-0144331.
 PR 19-JUL-1999; 990US-0144332.
 PR 19-JUL-1999; 990US-0144334.
 PR 19-JUL-1999; 990US-0144335.
 PR 20-JUL-1999; 990US-0144352.
 PR 20-JUL-1999; 990US-0144632.
 PR 20-JUL-1999; 990US-0144884.
 PR 21-JUL-1999; 990US-0144814.
 PR 21-JUL-1999; 990US-0145086.
 PR 21-JUL-1999; 990US-0145088.
 PR 22-JUL-1999; 990US-0145085.
 PR 22-JUL-1999; 990US-0145087.
 PR 22-JUL-1999; 990US-0145089.
 PR 22-JUL-1999; 990US-0145192.
 PR 23-JUL-1999; 990US-0145145.
 PR 23-JUL-1999; 990US-0145218.
 PR 23-JUL-1999; 990US-0145224.
 PR 26-JUL-1999; 990US-0145276.
 PR 27-JUL-1999; 990US-0145913.
 PR 27-JUL-1999; 990US-0145918.
 PR 27-JUL-1999; 990US-0145919.
 PR 28-JUL-1999; 990US-0145951.
 PR 02-AUG-1999; 990US-0146386.
 PR 02-AUG-1999; 990US-0146388.
 PR 02-AUG-1999; 990US-0146389.
 PR 03-AUG-1999; 990US-0147038.
 PR 04-AUG-1999; 990US-0147204.
 PR 04-AUG-1999; 990US-0147302.
 PR 05-AUG-1999; 990US-0147192.
 PR 05-AUG-1999; 990US-0147260.
 PR 06-AUG-1999; 990US-0147303.
 PR 06-AUG-1999; 990US-0147416.
 PR 09-AUG-1999; 990US-0147493.
 PR 09-AUG-1999; 990US-0147935.
 PR 10-AUG-1999; 990US-0148171.

PR	11-AUG-1999;	99US-0148319.	Oy	61 RHOADICHAYOLRKGGCLKDENILIVEMYDDIAFSENPPRCVILINKPGEDVYKGVDY 120
PR	12-AUG-1999;	99US-0148341.	Db	80 RHOADICHAYOLRKGLKDENIIVEMYDDIAFSENPPRCVILINKPGEDVYKGVPDY 139
PR	13-AUG-1999;	99US-0148565.	Qy	121 *REAVVNQNFNVLGNESGYTGGNGKVKGPNDFIYADHGAPGLIAMPGDEMYA 180
PR	16-AUG-1999;	99US-0148684.	Db	140 *REAVVNQNFNVLGNESGYTGGNGKVKGPNDFIYADHGAPGLIAMPGDEMYA 199
PR	17-AUG-1999;	99US-0149368.	Qy	121 *REAVVNQNFNVLGNESGYTGGNGKVKGPNDFIYADHGAPGLIAMPGDEMYA 180
PR	18-AUG-1999;	99US-0149415.	Db	140 *REAVVNQNFNVLGNESGYTGGNGKVKGPNDFIYADHGAPGLIAMPGDEMYA 199
PR	20-AUG-1999;	99US-0149426.	Qy	121 *REAVVNQNFNVLGNESGYTGGNGKVKGPNDFIYADHGAPGLIAMPGDEMYA 180
PR	20-AUG-1999;	99US-0149723.	Db	140 *REAVVNQNFNVLGNESGYTGGNGKVKGPNDFIYADHGAPGLIAMPGDEMYA 199
PR	20-AUG-1999;	99US-0149923.	Qy	181 KDFNEYLEKMKRKYKVNKMIVYEACESGSMEGILKKLNLYTAANNSKESSWGVYCP 240
PR	23-AUG-1999;	99US-0149902.	Db	200 KDFNEYLEKMKRKYKVNKMIVYEACESGSMEGILKKLNLYTAANNSKESSWGVYCP 259
PR	25-AUG-1999;	99US-0149930.	Qy	241 ESYPPPESEIGTCLGDFESISWLEDLHMSKETLEQQYHVYKRRVGSDVPIETSHVCRF 300
PR	26-AUG-1999;	99US-0150566.	Db	260 ESYPPPESEIGTCLGDFESISWLEDLHMSKETLEQQYHVYKRRVGSDVPIETSHVCRF 319
PR	27-AUG-1999;	99US-0150884.	Qy	301 GTEKMLKDYLSSYIGNPNENDNFTTESSESPISGLNPNRDPFLYLQRKTKOPAGS 360
PR	27-AUG-1999;	99US-0151065.	Db	320 GTEKMLKDYLSSYIGNPNENDNFTTESSESPISGLNPNRDPFLYLQRKTKOPAGS 379
PR	27-AUG-1999;	99US-0151066.	Qy	361 LESKEAQKKLDEKNRKQDOSTDILLESVKQTVNLNLTSRTTGOPLYWDWDCFKT 420
PR	27-AUG-1999;	99US-0151080.	Db	380 LESKEAQKKLDEKNRKQDOSTDILLESVKQTVNLNLTSRTTGOPLYWDWDCFKT 439
PR	30-AUG-1999;	99US-0151303.	Qy	421 LYNSFKNHGATVYHLKTYALANICNMGVDKOTVSAIEQACSM 466
PR	31-AUG-1999;	99US-0151438.	Db	440 LYNSFKNHGATVYHLKTYALANICNMGVDKOTVSAIEQACSM 485
PR	01-SEP-1999;	99US-0151930.		
PR	07-SEP-1999;	99US-0152363.		
PR	10-SEP-1999;	99US-0153070.		
PR	13-SEP-1999;	99US-0153758.		
PR	16-SEP-1999;	99US-0154018.		
PR	16-SEP-1999;	99US-0154039.		
PR	22-SEP-1999;	99US-0155119.		
PR	23-SEP-1999;	99US-0155486.		
PR	24-SEP-1999;	99US-0155659.		
PR	28-SEP-1999;	99US-0156438.		
PR	29-SEP-1999;	99US-0156596.		
PR	04-OCT-1999;	99US-0157117.		
PR	05-OCT-1999;	99US-0157753.		
PR	06-OCT-1999;	99US-0157855.		
PR	07-OCT-1999;	99US-0158029.		
PR	08-OCT-1999;	99US-0158232.		
PR	12-OCT-1999;	99US-0158365.		
PR	13-OCT-1999;	99US-0159293.		
PR	13-OCT-1999;	99US-0159294.		
PR	14-OCT-1999;	99US-0159225.		
PR	14-OCT-1999;	99US-0159329.		
PR	14-OCT-1999;	99US-0159330.		
PR	14-OCT-1999;	99US-0159331.		
PR	14-OCT-1999;	99US-0159637.		
PR	14-OCT-1999;	99US-0159638.		
PR	21-OCT-1999;	99US-0159584.		
PR	21-OCT-1999;	99US-0160741.		
PR	21-OCT-1999;	99US-0160767.		
PR	21-OCT-1999;	99US-0160778.		
PR	21-OCT-1999;	99US-0160779.		
PR	21-OCT-1999;	99US-0160814.		
PR	22-OCT-1999;	99US-0160915.		
PR	22-OCT-1999;	99US-0160981.		
PR	25-OCT-1999;	99US-0160999.		
PR	25-OCT-1999;	99US-0161404.		
PR	25-OCT-1999;	99US-0161405.		
PR	26-OCT-1999;	99US-0161359.		
PR	26-OCT-1999;	99US-0161360.		
PR	28-OCT-1999;	99US-0161361.		
PR	28-OCT-1999;	99US-0161920.		
PR	28-OCT-1999;	99US-0161992.		
PR	29-OCT-1999;	99US-0162142.		
Query Match	Best Local Similarity	99.78;	Score	2446;
Matches	465; Conservative	99.88;	Pred. No.	3.1e-209;
		0;	Mismatches	1;
		0;	Indels	0;
		0;	Gaps	0;
Qy	1 MSSPLGHFOILVLFHALLIFSAERSKTOOLLNDNVESSDKSAKTRWAVLGASNEYYNY 60		PR	04-MAY-1999;
Db	20 MSSPLGHFOILVLFHALLIFSAERSKTOOLLNDNVESSDKSAKTRWAVLGASNEYYNY 79		PR	05-MAY-1999;
			PR	06-MAY-1999;
			PR	07-MAY-1999;
			PR	08-MAY-1999;
			PR	09-MAY-1999;
			PR	10-MAY-1999;
			PR	11-MAY-1999;
			PR	12-MAY-1999;
			PR	13-MAY-1999;
			PR	14-MAY-1999;
			PR	15-MAY-1999;
			PR	16-MAY-1999;
			PR	17-MAY-1999;
			PR	18-MAY-1999;
			PR	19-MAY-1999;
			PR	20-MAY-1999;
			PR	21-MAY-1999;
			PR	22-MAY-1999;
			PR	23-MAY-1999;
			PR	24-MAY-1999;
			PR	25-MAY-1999;
			PR	26-MAY-1999;
			PR	27-MAY-1999;
			PR	28-MAY-1999;
			PR	29-MAY-1999;
			PR	30-MAY-1999;
			PR	31-MAY-1999;
			PR	01-JUN-1999;
			PR	02-JUN-1999;
			PR	03-JUN-1999;
			PR	04-JUN-1999;
			PR	05-JUN-1999;
			PR	06-JUN-1999;
			PR	07-JUN-1999;
			PR	08-JUN-1999;
			PR	09-JUN-1999;
			PR	10-JUN-1999;
			PR	11-JUN-1999;
			PR	12-JUN-1999;
			PR	13-JUN-1999;
			PR	14-JUN-1999;
			PR	15-JUN-1999;
			PR	16-JUN-1999;
			PR	17-JUN-1999;
			PR	18-JUN-1999;
			PR	19-JUN-1999;
			PR	20-JUN-1999;
			PR	21-JUN-1999;
			PR	22-JUN-1999;
			PR	23-JUN-1999;
			PR	24-JUN-1999;
			PR	25-JUN-1999;
			PR	26-JUN-1999;
			PR	27-JUN-1999;
			PR	28-JUN-1999;
			PR	29-JUN-1999;
			PR	30-JUN-1999;
			PR	31-JUN-1999;
			PR	01-JUL-1999;
			PR	02-JUL-1999;
			PR	03-JUL-1999;
			PR	04-JUL-1999;
			PR	05-JUL-1999;
			PR	06-JUL-1999;
			PR	07-JUL-1999;
			PR	08-JUL-1999;
			PR	09-JUL-1999;
			PR	10-JUL-1999;
			PR	11-JUL-1999;
			PR	12-JUL-1999;
			PR	13-JUL-1999;
			PR	14-JUL-1999;
			PR	15-JUL-1999;
			PR	16-JUL-1999;
			PR	17-JUL-1999;
			PR	18-JUL-1999;
			PR	19-JUL-1999;
			PR	20-JUL-1999;
			PR	21-JUL-1999;
			PR	22-JUL-1999;
			PR	23-JUL-1999;
			PR	24-JUL-1999;
			PR	25-JUL-1999;
			PR	26-JUL-1999;
			PR	27-JUL-1999;
			PR	28-JUL-1999;
			PR	29-JUL-1999;
			PR	30-JUL-1999;
			PR	31-JUL-1999;
			PR	01-AUG-1999;
			PR	02-AUG-1999;
			PR	03-AUG-1999;
			PR	04-AUG-1999;
			PR	05-AUG-1999;
			PR	06-AUG-1999;
			PR	07-AUG-1999;
			PR	08-AUG-1999;
			PR	09-AUG-1999;
			PR	10-AUG-1999;
			PR	11-AUG-1999;
			PR	12-AUG-1999;
			PR	13-AUG-1999;
			PR	14-AUG-1999;
			PR	15-AUG-1999;
			PR	16-AUG-1999;
			PR	17-AUG-1999;
			PR	18-AUG-1999;
			PR	19-AUG-1999;
			PR	20-AUG-1999;
			PR	21-AUG-1999;
			PR	22-AUG-1999;
			PR	23-AUG-1999;
			PR	24-AUG-1999;
			PR	25-AUG-1999;
			PR	26-AUG-1999;
			PR	27-AUG-1999;
			PR	28-AUG-1999;
			PR	29-AUG-1999;
			PR	30-AUG-1999;
			PR	31-AUG-1999;
			PR	01-SEP-1999;
			PR	02-SEP-1999;
			PR	03-SEP-1999;
			PR	04-SEP-1999;
			PR	05-SEP-1999;
			PR	06-SEP-1999;
			PR	07-SEP-1999;
			PR	08-SEP-1999;
			PR	09-SEP-1999;
			PR	10-SEP-1999;
			PR	11-SEP-1999;
			PR	12-SEP-1999;
			PR	13-SEP-1999;
			PR	14-SEP-1999;
			PR	15-SEP-1999;
			PR	16-SEP-1999;
			PR	17-SEP-1999;
			PR	18-SEP-1999;
			PR	19-SEP-1999;
			PR	20-SEP-1999;
			PR	21-SEP-1999;
			PR	22-SEP-1999;
			PR	23-SEP-1999;
			PR	24-SEP-1999;
			PR	25-SEP-1999;
			PR	26-SEP-1999;
			PR	27-SEP-1999;
			PR	28-SEP-1999;
			PR	29-SEP-1999;
			PR	30-SEP-1999;
			PR	31-SEP-1999;
			PR	01-OCT-1999;
			PR	02-OCT-1999;
			PR	03-OCT-1999;
			PR	04-OCT-1999;
			PR	05-OCT-1999;
			PR	06-OCT-1999;
			PR	07-OCT-1999;
			PR	08-OCT-1999;
			PR	09-OCT-1999;
			PR	10-OCT-1999;
			PR	11-OCT-1999;
			PR	12-OCT-1999;
			PR	13-OCT-1999;
			PR	14-OCT-1999;
			PR	15-OCT-1999;
			PR	16-OCT-1999;
			PR	17-OCT-1999;
			PR	18-OCT-1999;
			PR	19-OCT-1999;
			PR	20-OCT-1999;
			PR	21-OCT-1999;
			PR	22-OCT-1999;
			PR	23-OCT-1999;
			PR	24-OCT-1999;
			PR	25-OCT-1999;
			PR	26-OCT-1999;
			PR	27-OCT-1999;
			PR	28-OCT-1999;
			PR	29-OCT-1999;
			PR	30-OCT-1999;
			PR	31-OCT-1999;
			PR	01-NOV-1999;
			PR	02-NOV-1999;
			PR	03-NOV-1999;
			PR	04-NOV-1999;
			PR	05-NOV-1999;
			PR	06-NOV-1999;
			PR	07-NOV-1999;
			PR	08-NOV-1999;
			PR	09-NOV-1999;
			PR	10-NOV-1999;
			PR	11-NOV-1999;
			PR	12-NOV-1999;
			PR	13-NOV-1999;
			PR	14-NOV-1999;

PR	07 MAY-1999;	99US-0132863;	23-JUL-1999;	99US-0145145;
PR	14 MAY-1999;	99US-0134256;	23-JUL-1999;	99US-0145218;
PR	14 MAY-1999;	99US-0134218;	23-JUL-1999;	99US-0145224;
PR	14 MAY-1999;	99US-0134219;	26-JUL-1999;	99US-0145276;
PR	21 MAY-1999;	99US-0134221;	27-JUL-1999;	99US-0145318;
PR	14 MAY-1999;	99US-0134370;	27-JUL-1999;	99US-0145318;
PR	18 MAY-1999;	99US-0134768;	27-JUL-1999;	99US-0145319;
PR	19 MAY-1999;	99US-0134941;	28-JUL-1999;	99US-0145351;
PR	20 MAY-1999;	99US-0135124;	02-AUG-1999;	99US-0146386;
PR	21 MAY-1999;	99US-0135353;	02-AUG-1999;	99US-0146388;
PR	25 MAY-1999;	99US-0135629;	02-AUG-1999;	99US-0146387;
PR	25 MAY-1999;	99US-0136021;	03-AUG-1999;	99US-0147038;
PR	27 MAY-1999;	99US-0136392;	04-AUG-1999;	99US-0147204;
PR	28 MAY-1999;	99US-0136782;	04-AUG-1999;	99US-0147204;
PR	01 JUN-1999;	99US-0137222;	05-AUG-1999;	99US-0147302;
PR	03 JUN-1999;	99US-0137528;	05-AUG-1999;	99US-0147302;
PR	04 JUN-1999;	99US-0137502;	05-AUG-1999;	99US-0147303;
PR	07 JUN-1999;	99US-0137724;	06-AUG-1999;	99US-0147303;
PR	08 JUN-1999;	99US-0138094;	06-AUG-1999;	99US-0147303;
PR	10 JUN-1999;	99US-0138540;	07-AUG-1999;	99US-0147303;
PR	14 JUN-1999;	99US-0138847;	07-AUG-1999;	99US-0147303;
PR	18 JUN-1999;	99US-0139119;	11-AUG-1999;	99US-0148319;
PR	18 JUN-1999;	99US-0139452;	12-AUG-1999;	99US-0148341;
PR	16 JUN-1999;	99US-0139453;	13-AUG-1999;	99US-0148341;
PR	17 JUN-1999;	99US-0139492;	13-AUG-1999;	99US-0148341;
PR	18 JUN-1999;	99US-0139454;	16-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139455;	10-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139455;	10-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139456;	11-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139457;	11-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139458;	12-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139459;	13-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139460;	16-AUG-1999;	99US-0148884;
PR	18 JUN-1999;	99US-0139461;	10-AUG-1999;	99US-0149002;
PR	18 JUN-1999;	99US-0139462;	10-AUG-1999;	99US-0149002;
PR	18 JUN-1999;	99US-0139463;	11-AUG-1999;	99US-0149175;
PR	18 JUN-1999;	99US-0139750;	20-AUG-1999;	99US-0149722;
PR	18 JUN-1999;	99US-0139817;	20-AUG-1999;	99US-0149722;
PR	22 JUN-1999;	99US-0139899;	23-AUG-1999;	99US-0149929;
PR	23 JUN-1999;	99US-0140353;	23-AUG-1999;	99US-0149929;
PR	01 JUL-1999;	99US-0141842;	17-AUG-1999;	99US-0149929;
PR	01 JUL-1999;	99US-0142154;	18-AUG-1999;	99US-0149929;
PR	24 JUN-1999;	99US-0140695;	25-AUG-1999;	99US-0149929;
PR	28 JUN-1999;	99US-0139763;	25-AUG-1999;	99US-0149929;
PR	29 JUN-1999;	99US-0140991;	30-AUG-1999;	99US-0150665;
PR	30 JUN-1999;	99US-0141287;	30-AUG-1999;	99US-0150665;
PR	01 JUL-1999;	99US-0141842;	31-AUG-1999;	99US-0151030;
PR	01 JUL-1999;	99US-0142154;	01-SEP-1999;	99US-0151030;
PR	02 JUL-1999;	99US-0142055;	02-SEP-1999;	99US-0151263;
PR	06 JUL-1999;	99US-0140823;	07-AUG-1999;	99US-0151066;
PR	08 JUL-1999;	99US-0142803;	27-AUG-1999;	99US-0151080;
PR	09 JUL-1999;	99US-0140991;	30-AUG-1999;	99US-0151093;
PR	12 JUL-1999;	99US-0142920;	15-SEP-1999;	99US-0151378;
PR	12 JUL-1999;	99US-0142977;	31-AUG-1999;	99US-0151438;
PR	19 JUL-1999;	99US-0143542;	16-SEP-1999;	99US-0150656;
PR	19 JUL-1999;	99US-0143542;	26-AUG-1999;	99US-0150884;
PR	14 JUL-1999;	99US-0143624;	27-AUG-1999;	99US-0149722;
PR	15 JUL-1999;	99US-0144005;	07-OCT-1999;	99US-0153070;
PR	16 JUL-1999;	99US-0144085;	24-SEP-1999;	99US-0153070;
PR	16 JUL-1999;	99US-0144086;	24-SEP-1999;	99US-0153070;
PR	19 JUL-1999;	99US-0143252;	08-OCT-1999;	99US-0158332;
PR	20 JUL-1999;	99US-0143522;	08-OCT-1999;	99US-0158332;
PR	20 JUL-1999;	99US-0146322;	12-OCT-1999;	99US-0158332;
PR	20 JUL-1999;	99US-0146322;	13-OCT-1999;	99US-0158332;
PR	21 JUL-1999;	99US-0144333;	13-OCT-1999;	99US-0158332;
PR	21 JUL-1999;	99US-0144334;	14-OCT-1999;	99US-0158332;
PR	19 JUL-1999;	99US-0144085;	07-OCT-1999;	99US-0158332;
PR	19 JUL-1999;	99US-0144086;	07-OCT-1999;	99US-0158332;
PR	19 JUL-1999;	99US-0144335;	08-OCT-1999;	99US-0158332;
PR	22 JUL-1999;	99US-0144335;	14-OCT-1999;	99US-0159331;
PR	22 JUL-1999;	99US-0145087;	14-OCT-1999;	99US-0159331;
PR	22 JUL-1999;	99US-0145192;	18-OCT-1999;	99US-0159384;
PR	21 JUL-1999;	99US-0145086;	21-OCT-1999;	99US-0160741;
PR	22 JUL-1999;	99US-0145085;	21-OCT-1999;	99US-0160741;
PR	22 JUL-1999;	99US-0145087;	21-OCT-1999;	99US-0160748;
PR	22 JUL-1999;	99US-0145192;	21-OCT-1999;	99US-0160811;

Query Match 81.9%; Score 2009; DB 21; Length 380;
 Best Local Similarity 100.0%; Pred. No. 2.1e-10; Indels 0; Gaps 0;
 Matches 380; Conservative 0; Mismatches 0; Insertions 0;

Qy 87 MYDDIAFSENPRGVLINKPDGEDVYKGVPKDTKEAVNVQNYNVLLGNESCVTGNGC 146
 Db 1 MYDDIAFSENPRGVLINKPDGEDVYKGVPKDTKEAVNVQNYNVLLGNESCVTGNG 60

Qy 147 KVKGENDNIFTYADHAGFLIAMPTGDEVMAKDFNEVLEKHKKRKYNNKAVIYVAC 206
 Db 61 KVKGENDNIFTYADHAGFLIAMPTGDEVMAKDFNEVLEKHKKRKYNNKAVIYVAC 120

Qy 207 ESSGMFEGILRKLNLTAYTAANNSKESWGVYCPESSYPPPESEGTCLGDTFSLSWLED 266
 Db 121 ESSGMFEGILRKLNLTAYTAANNSKESWGVYCPESSYPPPESEGTCLGDTFSLSWEDS 180

Qy 267 DHDMSKETLHQYHVVKRYGSVPETSHVCRPSTERMLKDLYSSYGRNPENDNFT 326
 Db 181 DHDMSKETLHQYHVVKRYGSVPETSHVCRPSTERMLKDLYSSYGRNPENDNFT 240

Qy 327 ESSPTNSGLVNPNIDIPLYLORKTOKAPMGSLSEKAQKKLDEKNHRKOIDSTID 386
 Db 241 ESSPTNSGLVNPNIDIPLYLQRKIQKAPMGSLSEKAQKKLDEKNHRKQDQSITD 300

Qy 387 ILLRSYKQTNVNLNLTSTRTGQPLVDDWDCFKTLYNSFKHCCATVHYGLKTGALANI 446
 Db 301 ILLRSYKQTNVNLNLTSTRTGQPLVDDWDCFKTLYNSFKHCCATVHYGLKTGALANI 360

Qy 447 CNMGVDVKQTVSAIEACSM 466
 Db 361 CNMGVDVKQTVSAIEACSM 380

RESULT 5
 ID AAC12662 standard; Protein: 380 AA.
 XX
 AC AAC12662;
 XX
 DT 17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 11862.

XX Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 XX termination sequence.

Arabidopsis thaliana.

XX EP1033405-A2.
 PN
 XX
 PD 05-SEP-2000.
 XX
 PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121205.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.

99US-0160980.
 99US-0160981.
 99US-0160989.
 99US-0161404.
 99US-0161405.
 99US-0161406.
 99US-0161359.
 99US-0161360.
 99US-0161361.
 99US-0161362.
 99US-0161920.
 99US-0161922.
 99US-0161923.
 99US-0162142.

PR 22-OCT-1999; 99US-0126264.
 PR 22-OCT-1999; 99US-0126765.
 PR 22-OCT-1999; 99US-0127492.
 PR 25-OCT-1999; 99US-0128244.
 PR 25-OCT-1999; 99US-0128714.
 PR 26-OCT-1999; 99US-0129845.
 PR 26-OCT-1999; 99US-0130077.
 PR 26-OCT-1999; 99US-0130459.
 PR 28-OCT-1999; 99US-0130510.
 PR 28-OCT-1999; 99US-0130831.
 PR 28-OCT-1999; 99US-0131449.
 PR 28-OCT-1999; 99US-0132048.
 PR 29-OCT-1999; 99US-0132407.
 PR 29-OCT-1999; 99US-0132444.
 PR 05-MAY-1999; 99US-0132455.
 PR 05-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132497.
 PR 07-MAY-1999; 99US-0132853.
 PR 11-MAY-1999; 99US-0134216.
 PR 14-MAY-1999; 99US-0134248.
 PR 14-MAY-1999; 99US-0134239.
 PR 14-MAY-1999; 99US-0134270.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135134.
 PR 21-MAY-1999; 99US-0135333.
 PR 24-MAY-1999; 99US-0135639.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136312.
 PR 01-JUN-1999; 99US-0136712.
 PR 03-JUN-1999; 99US-0137538.
 PR 04-JUN-1999; 99US-0137562.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138044.
 PR 10-JUN-1999; 99US-0138540.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139412.
 PR 16-JUN-1999; 99US-0139413.
 PR 17-JUN-1999; 99US-0139412.
 PR 18-JUN-1999; 99US-0139414.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139753.
 PR 18-JUN-1999; 99US-0139750.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 02-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142300.
 PR 09-JUL-1999; 99US-0142863.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143512.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.

PR	16-JUL-1999;	990US-0144086;	PR	08-OCT-1999;	990US-0158232.
PR	19-JUL-1999;	990US-0144325;	PR	12-OCT-1999;	990US-0158369.
PR	19-JUL-1999;	990US-0144331;	PR	13-OCT-1999;	990US-0159293.
PR	19-JUL-1999;	990US-0144332;	PR	13-OCT-1999;	990US-0159294.
PR	19-JUL-1999;	990US-0144333;	PR	13-OCT-1999;	990US-0159295.
PR	19-JUL-1999;	990US-0144334;	PR	14-OCT-1999;	990US-0159329.
PR	19-JUL-1999;	990US-0144335;	PR	14-OCT-1999;	990US-0159330.
PR	20-JUL-1999;	990US-0144335;	PR	14-OCT-1999;	990US-0159331.
PR	20-JUL-1999;	990US-0144632;	PR	14-OCT-1999;	990US-0159637.
PR	20-JUL-1999;	990US-0144884;	PR	14-OCT-1999;	990US-0159638.
PR	21-JUL-1999;	990US-0144814;	PR	18-OCT-1999;	990US-0159584.
PR	21-JUL-1999;	990US-0145086;	PR	21-OCT-1999;	990US-0160741.
PR	21-JUL-1999;	990US-0145088;	PR	21-OCT-1999;	990US-0160767.
PR	22-JUL-1999;	990US-0145085;	PR	21-OCT-1999;	990US-0160768.
PR	22-JUL-1999;	990US-0145087;	PR	21-OCT-1999;	990US-0160769.
PR	22-JUL-1999;	990US-0145089;	PR	21-OCT-1999;	990US-0160814.
PR	22-JUL-1999;	990US-0145192;	PR	21-OCT-1999;	990US-0160815.
PR	23-JUL-1999;	990US-0145145;	PR	22-OCT-1999;	990US-0160980.
PR	23-JUL-1999;	990US-0145218;	PR	22-OCT-1999;	990US-0160981.
PR	23-JUL-1999;	990US-0145224;	PR	22-OCT-1999;	990US-0160989.
PR	26-JUL-1999;	990US-0145276;	PR	25-OCT-1999;	990US-0161404.
PR	27-JUL-1999;	990US-0145913;	PR	25-OCT-1999;	990US-0161405.
PR	27-JUL-1999;	990US-0145918;	PR	25-OCT-1999;	990US-0161406.
PR	27-JUL-1999;	990US-0145919;	PR	26-OCT-1999;	990US-0161359.
PR	02-AUG-1999;	990US-0145951;	PR	26-OCT-1999;	990US-0161360.
PR	02-AUG-1999;	990US-0146386;	PR	26-OCT-1999;	990US-0161361.
PR	02-AUG-1999;	990US-0146388;	PR	28-OCT-1999;	990US-0161920.
PR	02-AUG-1999;	990US-0146389;	PR	28-OCT-1999;	990US-0161992.
PR	03-AUG-1999;	990US-0147038;	PR	28-OCT-1999;	990US-0161993.
PR	04-AUG-1999;	990US-0147204;	PR	29-OCT-1999;	990US-0162142.
PR	05-AUG-1999;	990US-0147192;			
PR	05-AUG-1999;	990US-0147260;	Query Match	81.5%	Score 2001;
PR	06-AUG-1999;	990US-0147303;	Best Local Similarity	99.7%	DB 21;
PR	06-AUG-1999;	990US-0147416;	Mismatches	0;	Pred. No. 1.1e-169.
PR	09-AUG-1999;	990US-0147493;	Indels	1;	
PR	09-AUG-1999;	990US-0147935;	Gaps	0;	
PR	10-AUG-1999;	990US-0147935;			
PR	11-AUG-1999;	990US-0148171;			
PR	12-AUG-1999;	990US-0148319;			
PR	13-AUG-1999;	990US-0148341;			
PR	13-AUG-1999;	990US-0148565;			
PR	16-AUG-1999;	990US-0148684;			
PR	17-AUG-1999;	990US-0148684;			
PR	18-AUG-1999;	990US-0149175;			
PR	20-AUG-1999;	990US-0149426;			
PR	20-AUG-1999;	990US-0149722;			
PR	20-AUG-1999;	990US-0149723;			
PR	23-AUG-1999;	990US-0149929;			
PR	23-AUG-1999;	990US-0151066;			
PR	25-AUG-1999;	990US-0151303;			
PR	26-AUG-1999;	990US-0150566;			
PR	01-SEP-1999;	990US-0150884;			
PR	07-SEP-1999;	990US-0151199;			
PR	13-SEP-1999;	990US-0151066;			
PR	13-SEP-1999;	990US-01513758;			
PR	15-SEP-1999;	990US-0154018;			
PR	16-SEP-1999;	990US-0154039;			
PR	20-SEP-1999;	990US-0154779;			
PR	22-SEP-1999;	990US-0155139;			
PR	24-SEP-1999;	990US-0155486;			
PR	28-SEP-1999;	990US-0155659;			
PR	29-SEP-1999;	990US-0156458;			
PR	04-OCT-1999;	990US-0157117;			
PR	05-OCT-1999;	990US-0157853;			
PR	06-OCT-1999;	990US-0158029;			
			Query Match	81.5%	Score 2001;
			Best Local Similarity	99.7%	DB 21;
			Mismatches	0;	Pred. No. 1.1e-169.
			Indels	1;	
			Gaps	0;	
QY	87 MYDDIAFSENPRPGVIIINKPDGSDVYKGPDKTDYKEAVNVNFVNLLNESGGVTGGNG 146		Matches	379;	Conservative
Db	1 MYDDIAFSENPRPGVIIINKPDGSDVYKGPDKTDYKEAVNVNFVNLLNESGGVTGGNG 60				
QY	147 KVVKSGPNDIIFYADIGAGPLAMPGEDEMAKDFNLEMKKKRKYVNMVYVEAC 206				
Db	61 KVVKSGPNDIIFYADIGAGPLAMPGEDEMAKDFNLEMKKKRKYVNMVYVEAC 120				
QY	207 ESGSMFEGILKKNLITYAVTAANSKESSWGVYCPESTPPPPSERGTCIGDTFSISWLED 266				
Db	121 ESGSMFEGILKKNLITYAVTAANSKESSWGVYCPESTPPPPSERGTCIGDTFSISWLED 180				
QY	267 DLHDMSKETLEQQHVKRVGSDPVTSHCRGTEKMLKDYLSSYIGRPNDNFTFT 326				
Db	181 DLHDMSKETLEQQHVKRVGSDPVTSHCRGTEKMLKDYLSSYIGRPNDNFTFT 240				
QY	327 ESSSPISNGLYNPRDPLLQRKTLQAPMSSLESKAQKTLDEKHNRKOIDQSITD 386				
Db	241 ESSSPISNGLYNPRDPLLQRKTLQAPMSSLESKAQKTLDEKHNRKOIDQSITD 300				
QY	387 ILRSVVKONVNLNLSTRTGQPLDVDCPKTLVNSFKHCGATVHYGLKYTGALANI 446				
Db	301 ILRSVVKONVNLNLSTRTGQPLDVDCPKTLVNSFKHCGATVHYGLKYTGALANI 360				
QY	447 CNMGVDVKQTVALEEQACSM 466				
Db	361 CNMGVDVKQTVALEEQACSM 380				
RESULT	6				
AAG52575	ID AAG52575 standard; Protein; 295 AA.				
AC	XX				
AC	AC				
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 66849.				
DE	Arabidopsis thaliana protein fragment (first entry)				

PR	XX	99US-0140352;
PR	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	KW	termination sequence.
PR	XX	<i>Arabidopsis thaliana</i> .
PR	OS	
PR	XX	
PR	PN	
PR	PN	EP1033405-A2.
PR	XX	
PR	PD	06-SEP-2000.
PR	XX	
PR	PF	25-FEB-2000; 20000EP-0301439.
PR	XX	
PR	PR	23-JUN-1999; 99US-0121825.
PR	PR	05-MAR-1999; 99US-0123180.
PR	PR	09-MAR-1999; 99US-0123548.
PR	PR	23-MAR-1999; 99US-0125788.
PR	PR	25-MAR-1999; 99US-0126264.
PR	PR	29-MAR-1999; 99US-0126785.
PR	PR	01-APR-1999; 99US-0127462.
PR	PR	06-APR-1999; 99US-0128234.
PR	PR	08-APR-1999; 99US-0128714.
PR	PR	16-APR-1999; 99US-0129845.
PR	PR	30-APR-1999; 99US-0130077.
PR	PR	21-APR-1999; 99US-0130419.
PR	PR	23-APR-1999; 99US-0130510.
PR	PR	23-APR-1999; 99US-0130891.
PR	PR	28-APR-1999; 99US-0131449.
PR	PR	30-APR-1999; 99US-0132048.
PR	PR	04-MAY-1999; 99US-0132407.
PR	PR	04-MAY-1999; 99US-0132484.
PR	PR	05-MAY-1999; 99US-0132485.
PR	PR	06-MAY-1999; 99US-0132486.
PR	PR	06-MAY-1999; 99US-0132487.
PR	PR	07-MAY-1999; 99US-0132863.
PR	PR	11-MAY-1999; 99US-0134236.
PR	PR	14-MAY-1999; 99US-0134218.
PR	PR	14-MAY-1999; 99US-0134219.
PR	PR	14-MAY-1999; 99US-0134221.
PR	PR	18-MAY-1999; 99US-0134370.
PR	PR	19-MAY-1999; 99US-0134941.
PR	PR	20-MAY-1999; 99US-0135124.
PR	PR	21-MAY-1999; 99US-0135353.
PR	PR	24-MAY-1999; 99US-0135629.
PR	PR	25-MAY-1999; 99US-0136021.
PR	PR	27-MAY-1999; 99US-0136392.
PR	PR	28-MAY-1999; 99US-0136782.
PR	PR	01-JUN-1999; 99US-0137224.
PR	PR	03-JUN-1999; 99US-0137528.
PR	PR	04-JUN-1999; 99US-0137502.
PR	PR	07-JUN-1999; 99US-0137724.
PR	PR	08-JUN-1999; 99US-0138034.
PR	PR	10-JUN-1999; 99US-0138510.
PR	PR	14-JUN-1999; 99US-0138847.
PR	PR	16-JUN-1999; 99US-0139119.
PR	PR	16-JUN-1999; 99US-0139152.
PR	PR	17-JUN-1999; 99US-0139457.
PR	PR	18-JUN-1999; 99US-0139459.
PR	PR	18-JUN-1999; 99US-0139454.
PR	PR	18-JUN-1999; 99US-0139455.
PR	PR	18-JUN-1999; 99US-0139456.
PR	PR	18-JUN-1999; 99US-0139457.
PR	PR	18-JUN-1999; 99US-0139458.
PR	PR	18-JUN-1999; 99US-0139459.
PR	PR	18-JUN-1999; 99US-0139460.
PR	PR	18-JUN-1999; 99US-0139461.
PR	PR	18-JUN-1999; 99US-0139462.
PR	PR	18-JUN-1999; 99US-0139463.
PR	PR	18-JUN-1999; 99US-0139750.
PR	PR	21-JUN-1999; 99US-0139817.
PR	PR	22-JUN-1999; 99US-0139839.

Db	300 QRIDALIHWEEKRAPEGSPEKNAAQKQVVEVMSHRMHDNGVELIGKLFGIEKGPKV	359		PR	10-JUN-1999;
QY	401 LISTRIGQPLIDWCEFKTLVNSFRNHCCATVHGLKTGALLANICNNGVDFQTVSAI	460		PR	10-JUN-1999;
Db	360 LDVVRPAGMALVDWDCLKTMVTRFETYCGSLSQYGMKHMRSFANICNAGIKNQMDAS	419		PR	14-JUN-1999;
Qy	461 EQAC 464			PR	16-JUN-1999;
Db	420 AQC 423			PR	17-JUN-1999;
XX	AAC31187 standard; Protein: 490 AA.			PR	18-JUN-1999;
AC	AAC31187			PR	18-JUN-1999;
XX	17-OCT-2000 (first entry)			PR	18-JUN-1999;
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 37411.			PR	21-JUN-1999;
XX	Protein identification; signal transduction pathway; metabolic pathway;			PR	22-JUN-1999;
KW	KW identification assay; genetic mapping; gene expression control; promoter;			PR	23-JUN-1999;
KW	KW termination sequence.			PR	23-JUN-1999;
XX	Arabidopsis thaliana.			PR	24-JUN-1999;
XX	EP1033405-A2.			PR	28-JUN-1999;
PD	06-SEP-2000.			PR	29-JUN-1999;
XX	25-FEB-2000; 2000EP-0301439.			PR	30-JUN-1999;
PF				PR	01-JUL-1999;
XX	25-FEB-1999; 9905-012185.			PR	02-JUL-1999;
PR	09-MAR-1999; 9905-0123180.			PR	03-JUL-1999;
PR	09-MAR-1999; 9905-0123548.			PR	04-JUL-1999;
PR	23-MAR-1999; 9905-0125788.			PR	05-JUL-1999;
PR	25-MAR-1999; 9905-0126264.			PR	06-JUL-1999;
PR	29-MAR-1999; 9905-0126755.			PR	07-JUL-1999;
PR	01-APR-1999; 9905-0127462.			PR	08-JUL-1999;
PR	06-APR-1999; 9905-0128234.			PR	09-JUL-1999;
PR	08-APR-1999; 9905-0128714.			PR	10-JUL-1999;
PR	16-APR-1999; 9905-0129845.			PR	11-JUL-1999;
PR	19-APR-1999; 9905-0130077.			PR	12-JUL-1999;
PR	21-APR-1999; 9905-0130449.			PR	13-JUL-1999;
PR	23-APR-1999; 9905-0130510.			PR	14-JUL-1999;
PR	28-APR-1999; 9905-0130891.			PR	15-JUL-1999;
PR	28-APR-1999; 9905-0131449.			PR	16-JUL-1999;
PR	30-APR-1999; 9905-0132048.			PR	17-JUL-1999;
PR	30-APR-1999; 9905-0132407.			PR	18-JUL-1999;
PR	04-MAY-1999; 9905-0132454.			PR	19-JUL-1999;
PR	05-MAY-1999; 9905-0132455.			PR	20-JUL-1999;
PR	06-MAY-1999; 9905-0132416.			PR	21-JUL-1999;
PR	07-MAY-1999; 9905-0132487.			PR	22-JUL-1999;
PR	11-MAY-1999; 9905-0132456.			PR	23-JUL-1999;
PR	14-MAY-1999; 9905-0132418.			PR	24-JUL-1999;
PR	14-MAY-1999; 9905-0134219.			PR	25-JUL-1999;
PR	14-MAY-1999; 9905-0134221.			PR	27-JUL-1999;
PR	18-MAY-1999; 9905-0134768.			PR	27-JUL-1999;
PR	19-MAY-1999; 9905-0134941.			PR	28-JUL-1999;
PR	20-MAY-1999; 9905-0135124.			PR	02-AUG-1999;
PR	21-MAY-1999; 9905-0135351.			PR	02-AUG-1999;
PR	24-MAY-1999; 9905-0135629.			PR	02-AUG-1999;
PR	25-MAY-1999; 9905-0136021.			PR	03-AUG-1999;
PR	27-MAY-1999; 9905-016332.			PR	04-AUG-1999;
PR	28-MAY-1999; 9905-016702.			PR	04-AUG-1999;
PR	01-JUN-1999; 9905-0137222.			PR	05-AUG-1999;
PR	03-JUN-1999; 9905-0137528.			PR	05-AUG-1999;
PR	04-JUN-1999; 9905-0137502.			PR	06-AUG-1999;
PR	07-JUN-1999; 9905-0137724.			PR	07-AUG-1999;
PR	08-JUN-1999; 9905-0138094.			PR	09-AUG-1999;

PR	09-AUG-1999;	990US-0147935;	Db	45 ENDDDSNSTRAVLVAAGSGTWNYRHQADICHAYQLLRKGLKEENITVFMYDDIANNY
PR	10-AUG-1999;	990US-0148171;	Db	46 ENPPGVTINKPGEDIVYKGVPKDTEKAVANQNFNYNLGNESGVGGNGKVKGSPND
PR	11-AUG-1999;	990US-0148319;	Qy	96 ENPPGVTINKPGEDIVYKGVPKDTEKAVANQNFNYNLGNESGVGGNGKVKGSPND
PR	12-AUG-1999;	990US-0148341;	Db	105 ENRPGTINSPKGVDYQGPKDTEGDDVNNDLNLFAVILGDKTAVGGSGKVVDGSPND
PR	13-AUG-1999;	990US-0148565;	Qy	156 NIFTYYADHGAPLAMPTGDEWAKDENVELEKMHKRKYNNKVVTCVACESCSMFEGI
PR	13-AUG-1999;	990US-0148684;	Db	165 HIFTFYSPHGGPGVILGMPTSPLYANDINDVLKKKHALGYTSLVFYIEACESGSIFEGL
PR	16-AUG-1999;	990US-0149368;	Qy	216 LKRNLLNIAVTAANNSKESWWGVCYCPESYPPPESEIGTCLGDTFSISWLEDSDLHDMSKET
PR	17-AUG-1999;	990US-0149426;	Db	225 LPPGLNIAVTTAASNAAEWSWWGTYCPGEEPSPPPEYEICLGLDLYSVANMEDSGMHNLOTE
PR	18-AUG-1999;	990US-0149426;	Qy	276 LEQOYHVVKRR ---VGSDVPETSHVCRGFTERMLKDYLSSYGRNPENDNFETE--SFS
PR	20-AUG-1999;	990US-0149722;	Db	330 LHQQEYELVRRTAPVGTSY --GSHVMQGDVGLSKDNFLDLYMTNPANDNPFADANSLK
PR	20-AUG-1999;	990US-0149723;	Qy	342 PP ---SRVYNQRDALVIIWEKYRKAPEGSARKEAQKQVLEAMSRLHLIDNSVILVGKI
PR	23-AUG-1999;	990US-0149929;	Db	390 SP-SNSGLVNPDIPLLYLQRKIQKAPMGSLESKEAQKQLLDEKNHRQDQSITDLRL
PR	23-AUG-1999;	990US-0149902;	Qy	391 SYKQTQNVNLNLTSSTRITCQPLVWDCEFKTLVNSFKNHCATVHYGLKTYGALANICNMG
PR	25-AUG-1999;	990US-0149930;	Db	450 LGFISRGPEVLNKVRSAGQPLVIDWNCLKQVRAFERHGSGLSQYGIKHMRSFANICNAG
PR	26-AUG-1999;	990US-0150566;	Qy	451 VDVKQTVSAIEQACS
PR	27-AUG-1999;	990US-0150884;	Db	465 PR ---I
PR	27-AUG-1999;	990US-0151065;	Qy	474 DB 460 IQMEQMEAASQACT
PR	30-AUG-1999;	990US-0151303;	RESULT 9	
PR	31-AUG-1999;	990US-0151438;	AAG31186	
PR	01-SEP-1999;	990US-0151930;	ID AAG31186 standard; Protein: 494 AA.	
PR	07-SEP-1999;	990US-0152363;	XX	
PR	10-SEP-1999;	990US-0153070;	AC	
PR	15-SEP-1999;	990US-0153758;	XX	
PR	15-SEP-1999;	990US-0154018;	XX	
PR	16-SEP-1999;	990US-0154779;	DT	17-OCT-2000 (first entry)
PR	20-SEP-1999;	990US-0154779;	XX	Arabidopsis thaliana protein fragment SEQ ID NO: 37410.
PR	22-SEP-1999;	990US-0155139;	XX	
PR	23-SEP-1999;	990US-0155486;	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR	28-SEP-1999;	990US-0155659;	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR	28-SEP-1999;	990US-0156458;	KW	termination sequence.
PR	29-SEP-1999;	990US-0156596;	XX	Arabidopsis thaliana.
PR	04-OCT-1999;	990US-0157117;	OS	
PR	05-OCT-1999;	990US-0157799;	EP1033405-A2.	
PR	06-OCT-1999;	990US-0157865;	XX	
PR	08-OCT-1999;	990US-0158029;	PN	
PR	12-OCT-1999;	990US-0158232;	XX	
PR	12-OCT-1999;	990US-0158369;	PD	06-SEP-2000.
PR	13-OCT-1999;	990US-0159293;	XX	
PR	13-OCT-1999;	990US-0159294;	PF	25-FEB-2000; 2000EP-0301439.
PR	13-OCT-1999;	990US-0159295;	XX	
PR	14-OCT-1999;	990US-0159329;	PR	25-FEB-1999;
PR	14-OCT-1999;	990US-0159330;	PR	05-MAR-1999;
PR	14-OCT-1999;	990US-0159331;	PR	09-MAR-1999;
PR	14-OCT-1999;	990US-0159637;	PR	23-MAR-1999;
PR	14-OCT-1999;	990US-0159638;	PR	29-MAR-1999;
PR	18-OCT-1999;	990US-0159584;	PR	01-APR-1999;
PR	21-OCT-1999;	990US-0160741;	PR	06-APR-1999;
PR	21-OCT-1999;	990US-0160767;	PR	08-APR-1999;
PR	21-OCT-1999;	990US-0160768;	PR	16-APR-1999;
PR	21-OCT-1999;	990US-0160814;	PR	19-APR-1999;
PR	21-OCT-1999;	990US-0160815;	PR	21-APR-1999;
PR	22-OCT-1999;	990US-0160980;	PR	23-APR-1999;
PR	22-OCT-1999;	990US-0160981;	PR	23-APR-1999;
PR	22-OCT-1999;	990US-0160989;	PR	28-APR-1999;
PR	25-OCT-1999;	990US-0161404;	PR	29-APR-1999;
PR	25-OCT-1999;	990US-0161405;	PR	01-APR-1999;
PR	25-OCT-1999;	990US-0161406;	PR	01-APR-1999;
PR	26-OCT-1999;	990US-0161359;	PR	01-APR-1999;
PR	26-OCT-1999;	990US-0161360;	PR	01-APR-1999;
PR	26-OCT-1999;	990US-0161361;	PR	01-APR-1999;
PR	28-OCT-1999;	990US-0161992;	PR	04-MAY-1999;
PR	28-OCT-1999;	990US-0161993;	PR	04-MAY-1999;
PR	29-OCT-1999;	990US-0162142;	PR	05-MAY-1999;
Query Match	50.8%;	Score 1246;	DB 21;	Length 490;
Best Local Similarity	53.6%;	Pred. No. 3..le-10;		
Matches 233;	Conservative 71;	Mismatches 121;	Indels 10;	Gaps 4;
OY	36 ESSDKSARKTRWAVLVAAGSGTWNYRHQADICHAYQILRKGGIKDENIVMFYDDIAFSS 95			

PR	06-MAY-1999;	99US-0132486.	PR	22-JUL-1999;	99US-0145089.
PR	06 MAY 1999;	99US-0132487.	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	99US-0132863.	PR	23-JUL-1999;	99US-0145195.
PR	11-MAY-1999;	99US-0134236.	PR	23-JUL-1999;	99US-0145238.
PR	14-MAY-1999;	99US-0134218.	PR	23-JUL-1999;	99US-0145244.
PR	14-MAY-1999;	99US-0134219.	PR	26-JUL-1999;	99US-0145246.
PR	14-MAY-1999;	99US-0134221.	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	99US-0134370.	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	99US-0134768.	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	99US-0134941.	PR	28-JUL-1999;	99US-0145921.
PR	20-MAY-1999;	99US-0135224.	PR	02-AUG-1999;	99US-0146186.
PR	21-MAY-1999;	99US-0135253.	PR	04-AUG-1999;	99US-0146188.
PR	24-MAY-1999;	99US-0135629.	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	99US-0136021.	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	99US-0136322.	PR	04-AUG-1999;	99US-0147042.
PR	28-MAY-1999;	99US-0136782.	PR	04-AUG-1999;	99US-0147042.
PR	01-JUN-1999;	99US-0137222.	PR	05 AUG 1999;	99US-0147192.
PR	03-JUN-1999;	99US-0137228.	PR	06-AUG-1999;	99US-0147303.
PR	04-JUN-1999;	99US-0137502.	PR	06-AUG-1999;	99US-0147316.
PR	08-JUN-1999;	99US-0137724.	PR	09-AUG-1999;	99US-0147793.
PR	10-JUN-1999;	99US-0138094.	PR	10-AUG-1999;	99US-0147835.
PR	10-JUN-1999;	99US-0138440.	PR	11-AUG-1999;	99US-0148111.
PR	11-JUN-1999;	99US-0138847.	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	99US-0139119.	PR	13-AUG-1999;	99US-0148565.
PR	16-JUN-1999;	99US-0139452.	PR	13-AUG-1999;	99US-0148564.
PR	18-JUN-1999;	99US-0139433.	PR	16-Aug 1999;	99US-0149158.
PR	18-JUN-1999;	99US-0139450.	PR	17-Aug 1999;	99US-0149155.
PR	18-JUN-1999;	99US-0139459.	PR	18-Aug 1999;	99US-0149416.
PR	18-JUN-1999;	99US-0139458.	PR	20-Aug 1999;	99US-0149722.
PR	18-JUN-1999;	99US-0139459.	PR	20-Aug 1999;	99US-0149733.
PR	18-JUN-1999;	99US-0139460.	PR	20-Aug 1999;	99US-0149939.
PR	18-JUN-1999;	99US-0139461.	PR	23-Aug 1999;	99US-0149932.
PR	18-JUN-1999;	99US-0139462.	PR	23-Aug 1999;	99US-0149940.
PR	18-JUN-1999;	99US-0139473.	PR	25-Aug 1999;	99US-0150566.
PR	18-JUN-1999;	99US-0139750.	PR	26-Aug 1999;	99US-0150884.
PR	18-JUN-1999;	99US-0139763.	PR	27-Aug 1999;	99US-0151065.
PR	21-JUN-1999;	99US-0139817.	PR	27-Aug 1999;	99US-0151066.
PR	22-JUN-1999;	99US-0139899.	PR	30-Aug 1999;	99US-0151080.
PR	30-JUN-1999;	99US-0141287.	PR	31-Aug 1999;	99US-0151103.
PR	01-JUL-1999;	99US-0140353.	PR	01-SEP-1999;	99US-0151338.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151330.
PR	24-JUN-1999;	99US-0140675.	PR	02-SEP-1999;	99US-0152163.
PR	28-JUN-1999;	99US-0140833.	PR	10-SEP-1999;	99US-0153010.
PR	29-JUN-1999;	99US-0140951.	PR	13-SEP-1999;	99US-0153158.
PR	01-JUL-1999;	99US-0141842.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0142154.	PR	16-SEP-1999;	99US-0154339.
PR	02-JUL-1999;	99US-0142055.	PR	20-SEP-1999;	99US-0154779.
PR	06-JUL-1999;	99US-0142850.	PR	22-SEP-1999;	99US-0155139.
PR	09-JUL-1999;	99US-0142930.	PR	23-SEP-1999;	99US-0155659.
PR	12-JUL-1999;	99US-0142977.	PR	24-SEP-1999;	99US-0156158.
PR	13-JUL-1999;	99US-0143342.	PR	28-SEP-1999;	99US-0156917.
PR	14-JUL-1999;	99US-0143544.	PR	04-OCT-1999;	99US-0157117.
PR	15-JUL-1999;	99US-0144005.	PR	05-OCT-1999;	99US-0157753.
PR	16-JUL-1999;	99US-0144055.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144086.	PR	07-OCT-1999;	99US-0158039.
PR	19-JUL-1999;	99US-0144226.	PR	08-OCT-1999;	99US-015832.
PR	19-JUL-1999;	99US-0144331.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159233.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159244.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159319.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159310.
PR	20-JUL-1999;	99US-0144352.	PR	14-OCT-1999;	99US-0159321.
PR	20-JUL-1999;	99US-0144353.	PR	14-OCT-1999;	99US-0159617.
PR	21-JUL-1999;	99US-0144834.	PR	14-OCT-1999;	99US-0159618.
PR	21-JUL-1999;	99US-0144834.	PR	18-OCT-1999;	99US-0159554.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160747.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.

PR	21-OCT-1999;	990US-0160814	PF	07-AUG-1992;	92JP-0231602.
PR	21-OCT-1999;	990US-0160815	XX		
PR	22-OCT-1999;	990US-0160980	PR	07-FEB-1992;	92JP-0056023.
PR	22-OCT-1999;	990US-0160981	XX		
PR	22-OCT-1999;	990US-0160989	PA	(SYUZU/)	SYUZO T.
PR	25-OCT-1999;	990US-0161404	XX		
PR	25-OCT-1999;	990US-0161405	DR	WPI: 1993-373587/47.	
PR	25-OCT-1999;	990US-0161406	N-PSDB;	AAQ50570.	
PR	26-OCT-1999;	990US-0161359	XX		
PR	26-OCT-1999;	990US-0161360	PT	New gene for encoding asparaginyl endo-peptidase - comprises 8	
PR	26-OCT-1999;	990US-0161361	PT	specified DNA sequences	
PR	28-OCT-1999;	990US-0161920	XX	Dislosure; Page 17-19; 35pp; Japanese.	
PR	28-OCT-1999;	990US-0161998	PS		
PR	29-OCT-1999;	990US-0161999	XX	A gene encoding asparaginyl endopeptidase is claimed.	
PR	29-OCT-1999;	990US-0162142	CC	B DNA sequences are given AAQ50570-9-661. The enzyme is a protease derived from a seed of Canavalia ensiformis which selectively hydrolyses C-terminus amide bond of L-asparagine residue (see AA43033 and AAR43041).	
Qy	36 ESSDKSAKGTRWAVLVAQSNEYYNRHODICHAYQILRKGGIURDENITIVMDDIAFSS 95	THE enzyme is useful for protein fragmentation and enzymatic peptide synthesis.	CC		
Db	49 ENPDDNSGTRWAVLVAQSSEYYNRHODICHAYQILRKGGIKEENITVFMDDIANNY 108	The primers given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the isolation of the fragments given in AAQ50569-75 and AAR430578-79, by PCR.	CC		
Qy	96 ENPRGVITINKPGDVKYGPVDYTKRPEAKVNVLLGNGESVTTGGNGKVKVSGPND 155	XX	Sequence 440 AA;		
Db	109 ENPRGTIINSPIRGKDVQGPVDYTGDDVNVNLNAFLVIGDITAVRGSGKVKVSGPND 168	Query Match 50.7%; Score 1245; DB 14; Length 440;			
Qy	156 NIFTYYADHGAPGLIAMPGDEMAKDFNEVLEKFMHKRKKYKMKVIVYACESGSMSMEGI 215	Best Local Similarity 54.4%; Pred. No. 3.5e-102;			
Db	169 HIFTFYSDHGGPVLGMPTSPYLANDVLKKHALGTYSKSLVFEACESGSLSIFEGL 228	Matches 230; Conservative 67; Mismatches 124; Indels 2; Gaps 2;			
Qy	216 LKRNLLNIVAYTAANSKESWSGVVCPSEPPBEPICLGDFTEISWLEDSDHDMSET 275	Qy 43 KGTTRWAVLVAQSNEYYNRHODICHAYQILRKGGIURDENITIVMDDIAFSSENPRPGV 102			
Db	229 LPBGLNLYATTASNAEESSWGTCTCGPEEPSPPBEYETCLGDLYSVANMEDSGMHNLOT 288	Db 2 EGTRWAVLVAQSNGYWNRYHQSDVCHAYQILRKGGIURDENITIVMDDIAFSENPRPGV 61			
Qy	276 LEQQYHVYKRR--VGSDPVTSHVCRCFGTEKMLKDLYSSYIGRNPDNFETTE--SFS 330	Qy 103 IINRPDGDDVYKGVCPKDYVTPKDYVTPVNNFLAATLGNKSAITGGSKGVNSGPNDHIFIYS 121			
Db	289 LRQOYELKRRATPAVGKSY -GSHVMOQIDVGTSKDNLDLYNTNPANDNFTEDANSLK 346	Db 62 ILNSPHGDDVYKGVCPKDYVGDYKDTVNNFLAATLGNKSAITGGSKGVNSGPNDHIFIYS 121			
Qy	331 SPISNSGLVNPDIPLYLQRKTOKAPMGSLSEKAQKKLDEQFKHRKQDOSITDILRL 390	Qy 163 DHCAGPLJAMPTCDEVYAKDNEVYLERKMKRKKYKAVIYACESESMSMEEGILKKNUNI 222			
Db	347 PP---SRVYNQRADLVFWEKVRKAPEGSARTEAQVQLRMASHRLHDNSVILVGKI 403	Db 122 DHGPGVJGMPSTSFLYASDLTEVLKKHAASTGTYKSLVFLAECEGSIFEGLPEGLN 181			
Qy	391 SVKQTQVNLLTISTRTRGQPLVWDCEKTLVNSFKHCGATYHYGLKYTGALANICNMG 450	Qy 223 YATAAANSKESSRGVYCPBSYPPPPSEITGCLGDTFSISWLSDSLDHMSKENEIQLK 282			
Db	404 LFCTSRGPEVLNKRSVRAQQLPVDWNCNLKNOVRAFERHCGSLSQYGTKHMRSFANTCNAG 463	Db 182 YATTAANAKRESSRGVYCPBSYPPPPSEITGCLGDTFSISWLSDSLDHMSKENEIQLK 341			
Qy	451 VDVKQTVAEQAQS 465	Qy 283 VK-RRVGSDVPETSHVCRCFGTEKMLKDLYSSYIGRNPDNFETTESSPSISNSGLVNP 341			
Db	464 IQEQMEEAASQACT 478	Db 242 VRARTINGNSIYGSHVMOQGDIELSKNNFLYLTGTNPANDNFTEVDK NSLYPPSKAVNQ 300			
Qy	RESULT 1 AAR43035 standard; Protein; 440 AA.	Qy 342 RDIPPLYLRQKIQKAPMSLESKEAQKLDEQFKHRKQDOSITDILRLSVKOTVNLL 401			
DT	24-MAY-1994 (first entry)	Db 301 RDAVLVHWDFKPKAPESARKAARKVOLEAMSHRHIDDSMKLIGKLFGTEKGPEVL 360			
XX	Asparaginylendopeptidase derived from clone 101.	Qy 402 TSTRRTGQPLVWDCEKTLVNSFKHCGATYHYGLKYTGALANICNMGVDYKOTVSAIE 461			
AC	AAR43035;	Db 361 SSVPAGQALVDDMDCLLKTVRFETHCGSLSQYGMHMRSPANFCNAGIQREOMAEASA 420			
XX	L-asparagine; primer; PCR; protein fragmentation;	Qy 462 QAC 464			
DE	Asparaginylendopeptidase; Canavalia ensiformis; seed;	Db 421 QAC 423			
XX	KW Asparaginylendopeptidase; Canavalia ensiformis; seed;	RESULT 11			
XX	KW L-asparagine; primer; PCR; protein fragmentation;	AAR43037			
OS	Canavalia ensiformis.	ID AAR43037 standard; Protein; 440 AA.			
XX	JP05276560-A.	XX DT 24-MAY-1994 (first entry)			
XX	26-OCT-1993.	XX DE Asparaginylendopeptidase derived from clone 103.			

XX	Asparaginylendopeptidase; Canavalia ensiformis; seed;
KW	L-asparagine; primer; PCR; Protein fragmentation;
KW	peptide synthesis.
XX	Canavalia ensiformis.
XX	DNase I standard; Protein: 440 AA.
OS	AAR43040
PN	AC
JP05276960-A.	AC
XX	XX
PD	XX
XX	XX
XX	XX
PF	XX
XX	XX
07-AUG-1992;	92JJP-0231602.
XX	XX
07-FEB-1992;	92JJP-0056023.
XX	XX
PA	PA
(SYUZU /) SYUZO T.	SYUZO T.
XX	XX
WPI;	1993-377587/47.
DR	PN
N-PSDB;	AAQ5072.
XX	XX
XX	XX
PT	PT
New gene for encoding asparaginyl endo-peptidase - comprises 8	PT
specified DNA sequences	PT
PT	PT
Disclosure: Page 20-22; 35pp; Japanese.	Page 27-29; 35pp; Japanese.
XX	XX
CC	A gene encoding asparaginylendopeptidase is claimed.
CC	8 DNA sequences are given (AAQ5059-66). The enzyme is a protease derived from a seed of Canavalia ensiformis which selectively hydrolyses C-terminal amide bond of L-asparagine residue (see AAR3033 and AAR43041).
CC	The enzyme is useful for protein fragmentation and enzymatic peptide synthesis.
CC	The primers given in AAQ50567-68, AAQ50576-77 and AAQ50583-90 were used in the isolation of the fragments given in AAQ50569-75 and AAQ50578-79, by PCR.
XX	XX
Sequence	440 AA:
Score	49.5%;
Best Local Similarity	53.3%;
Matches	226;
Conservative	65;
Mismatches	127;
Indels	6;
Gaps	3;
Db	INKPQGEDYKGVFKDYTKEAVNVQNFYVNLIGNESEGVGGNGKVYSGPNDNIFIYAD 163
Qy	INKPQGDYYGKVDYKGDYQDADNFYAVILNKNSALTGSGKVYSGPNDRIFFYSD 122
Db	63 INKPQGDYYGKVDYKGDYQDADNFYAVILNKNSALTGSGKVYSGPNDRIFFYSD 122
Qy	44 GTRWAVLVAGSNYYNYRHOADCCHAQVYLKGGLDENITYVYMDIAFSENPRPGV 103
Db	3 GTRWAVLVAGSNYYNYRHOADCCHAQVYLKGGLDENITYVYMDIAFSENPRPGV 103
Qy	104 INKPQGEDYKGVFKDYTKEAVNVQNFYVNLIGNESEGVGGNGKVYSGPNDNIFIYAD 163
Db	63 INKPQGDYYGKVDYKGDYQDADNFYAVILNKNSALTGSGKVYSGPNDRIFFYSD 122
Qy	44 GTRWAVLVAGSNYYNYRHOADCCHAQVYLKGGLDENITYVYMDIAFSENPRPGV 103
Db	3 GTRWAVLVAGSNYYNYRHOADCCHAQVYLKGGLDENITYVYMDIAFSENPRPGV 103
Qy	104 INKPQGEDYKGVFKDYTKEAVNVQNFYVNLIGNESEGVGGNGKVYSGPNDNIFIYAD 163
Db	63 INKPQGDYYGKVDYKGDYQDADNFYAVILNKNSALTGSGKVYSGPNDRIFFYSD 122
Qy	164 HGAPGLIAMPTEGYMDEVMARDENYELEKHHKRKNKMYTYVERACEGSMFEGTLKKUNIV 223
Db	1.23 HGPGLVLMGPAPPFLYASLIEVLLKKHASGTYSKLVFYLEAEESGTSFEGLPEDINTY 182
Qy	224 AVTAANSKESSWGGYCPPEYPPPESEIGRLGTFISWLBDSDHMSKEPLQQYHV 283
Db	183 ATTAANAEESWGGYCPGDPSPPEYTCGLGEYSAWMDSDRKMLRTETLHQOYELV 242
Qy	284 KRR-NGSDPTEHVCRGTEKMLKDYLSSYIGRNPDNETE--SFSSPSNSGLVN 340
Db	243 KERTINGSIHSHSHMQIDISDDVLFYLGTNPANDNFYDNLSP--SKAIS 299
Qy	341 PRDIPLYLQRKIQKAPMGSLSESRAQKRLDKRKOIDSITDLRLSVQTQTNVLNL 400
Db	300 QRVADLHWEKFRAPEGSTTRDAQOFLEMDSRKHIDSVKIGSLPFGIERPEV 359
Qy	401 LSTRTRGQLVDDDCFTLVLNSFKNHCGATHYGHLYKTGALANICNMGVYKORYSAI 460
Db	360 LNVRPGMHLVDCLRNNAVTRPETYCGSSLQXGKYMRSFANICNARINDQRADAS 419
Qy	461 EQAC 464
Db	420 AQAC 423

QY 339 ---VNPRDIPILLYLQRKIQKAPMGSLESKEAQKKLLDEKNHRQIDOSITDILRLSVKQT 395
 PR 07-JUN-1999; 99US-0137724.
 DB 295 MERVVNQRDAELLFWQNYQRSNHQPENRKTHILEQITETVKIRNHLDGSVENJGVULLYGGP 354
 PR 99US-0138094; 99US-0138540.
 QY 396 NVLNLLSTSTRTGQPLVDDWDFKTLVNSFKNHCATVHGKTYTGTALANICNMGVDVQK 455
 PR 07-JUN-1999; 99US-0138847.
 DB 355 KSSVLSVHSVRAPGLPLVDDWTCLKSMSVRFETHCGSLTQCMKMHMRAFGNVNCNSGV---- 410
 PR 10-JUN-1999; 99US-0139119.
 QY 456 TVSAIEQAC 464
 PR 14-MAY-1999; 99US-0139452.
 DB 411 SKASMEAAC 419
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 XX PR 18-JUN-1999; 99US-0139462.
 AC AAG31188 standard; Protein; 395 AA.
 PR 18-JUN-1999; 99US-0139463.
 XX DT 17-OCT-2000 (first entry)
 PR 18-JUN-1999; 99US-0139750.
 XX PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139811.
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 37412.
 PR 21-JUN-1999; 99US-0139899.
 XX KW Protein identification; signal transduction pathway; metabolic pathway;
 PR 23-JUN-1999; 99US-0140354.
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 PR 24-JUN-1999; 99US-0140695.
 KW termination sequence.
 PR 28-JUN-1999; 99US-0140823.
 XX PR 29-JUN-1999; 99US-0140991.
 OS PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PN EP1033405-A2.
 PR 01-JUL-1999; 99US-0142154.
 XX PD 06-SEP-2000.
 PR 02-JUL-1999; 99US-0142055.
 XX PD 06-JUL-1999; 99US-0142390.
 PR 06-JUL-1999; 99US-0142803.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 XX PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126783.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132449.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 06-MAY-1999; 99US-0132487.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0135353.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0136229.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.

PR	06-AUG-1999;	99US-0147416;	Qy	87 MYDDIAFSSENPRPGVIIINKPGEDEVYK&PVKDYTKAEVNQNQNFYINVILGNESGVGGNG 146
PR	09-AUG-1999;	99US-0147493;	Db	1 MYDDANNENPRPGTIIINSPHGKVYQGPVKDYGDDVNVDNLFAVILGDKTAVKGSG 60
PR	10-AUG-1999;	99US-0147935;	Qy	147 KVKKGSPNDNIFIYADHGAPGLIAMPTEVMAKDFNEVLEMKHKKRKYKVNIVYAC 206
PR	11-AUG-1999;	99US-0148319;	Db	61 KVVDGSPNDHIFIYSDHGGPKVGLMPSPLYANDLNDVLKKKHGTYKSILVYLEAC 120
PR	12-AUG-1999;	99US-0148341;	Qy	207 ESGSHFEGGLUKNNIUYAVTAANSKESSWGVCPSEIGTCGDTFSISWLEDS 266
PR	13-AUG-1999;	99US-0148565;	Db	121 ESGSFEGLPEGUNIYATTASNAEESSESSGTYCPEERSPPPEYETCGDLYSVAMMEDS 180
PR	13-AUG-1999;	99US-0148684;	Qy	267 DLHDMSKETLEQOQHVVKRR--VGSQVPETSHVCREFTEKMLKDQYLSYIGRNPDNDNF 323
PR	16-AUG-1999;	99US-0149369;	Db	181 GMHNHQETEHQOLVKTAPGYSY -GSHMQYGDVGISKDNLDLYMGTNPANDNF 238
PR	17-AUG-1999;	99US-0149175;	Qy	324 TETE--SFSSPISNSGLVNPRDIPLYLQRKIQKAPMGSLSESKEAQKLLDEENHRKOID 381
PR	18-AUG-1999;	99US-0149426;	Db	239 TFDADNSLRP---SRVNTORDADLVHEPEYKRAPESSARKEAQKVOLEANSRHLHD 295
PR	20-AUG-1999;	99US-0149722;	Qy	382 QSTIDILRSVKONVNVLNLTSTTGOPAUVDMDICFKTLVNSPKHNGATVHYGLKTG 441
PR	23-AUG-1999;	99US-0149902;	Db	296 NSVILVGKLFISRGPSPELNKVRSAQPLVDDWNCLKNQVRAFERHCCGSLSQYGKHM 355
PR	25-AUG-1999;	99US-0150566;	Qy	442 ALANICNMGEVDVQTVSAEQAC 465
PR	26-AUG-1999;	99US-0150884;	Db	356 SFANNCNAGTQMEEAQSQACT 379
PR	27-AUG-1999;	99US-0151065;		
PR	27-AUG-1999;	99US-0151066;		
PR	27-AUG-1999;	99US-0151080;		
PR	30-AUG-1999;	99US-0151303;		
PR	31-AUG-1999;	99US-0151438;		
PR	01-SEP-1999;	99US-0151930;		
PR	07-SEP-1999;	99US-0152363;		
PR	10-SEP-1999;	99US-0153070;		
PR	13-SEP-1999;	99US-0153758;		
PR	15-SEP-1999;	99US-0154018;		
PR	16-SEP-1999;	99US-0154039;		
PR	20-SEP-1999;	99US-0154779;		
PR	22-SEP-1999;	99US-0155139;		
PR	24-SEP-1999;	99US-0155659;		
PR	28-SEP-1999;	99US-0156458;		
PR	29-SEP-1999;	99US-0155516;		
PR	04-OCT-1999;	99US-0157117;		
PR	05-OCT-1999;	99US-0157753;		
PR	06-OCT-1999;	99US-0157865;		
PR	07-OCT-1999;	99US-0158029;		
PR	08-OCT-1999;	99US-0158222;		
PR	12-OCT-1999;	99US-0158369;		
PR	13-OCT-1999;	99US-0159293;		
PR	13-OCT-1999;	99US-0159295;		
PR	14-OCT-1999;	99US-0159322;		
PR	14-OCT-1999;	99US-0159330;		
PR	14-OCT-1999;	99US-0159331;		
PR	14-OCT-1999;	99US-0159631;		
PR	18-OCT-1999;	99US-0159638;		
PR	21-OCT-1999;	99US-0160741;		
PR	21-OCT-1999;	99US-0160767;		
PR	21-OCT-1999;	99US-0160768;		
PR	21-OCT-1999;	99US-0160770;		
PR	21-OCT-1999;	99US-0160814;		
PR	21-OCT-1999;	99US-0160815;		
PR	22-OCT-1999;	99US-0160980;		
PR	22-OCT-1999;	99US-0160981;		
PR	25-OCT-1999;	99US-0160989;		
PR	25-OCT-1999;	99US-0161404;		
PR	25-OCT-1999;	99US-0161445;		
PR	26-OCT-1999;	99US-0161406;		
PR	26-OCT-1999;	99US-0161359;		
PR	26-OCT-1999;	99US-0161360;		
PR	28-OCT-1999;	99US-0161361;		
PR	28-OCT-1999;	99US-0161920;		
PR	28-OCT-1999;	99US-0161933;		
PR	29-OCT-1999;	99US-0162142;		
	Query Match	41.7%	Score 1023;	DB 21;
	Best Local Similarity	50.3%	Pred. No. 1.9e-06;	Length 395;
	Matches 193;	Conservative	65;	Gaps 4;
			Mismatches 116;	Indels 10;
				PS XX
				DR N-PSDB: ANV44866.
				XX
				New isolated osteoclast inhibitor protein - used to develop products
				which can be used in the method of the invention. The method is
				for inhibiting osteoclastogenesis, and comprises administering to a ce
				an osteoclast (OC) inhibitor protein (OIP) composition in a vehicle
				where the composition inhibits production of OCs. The OIP polypeptides
				have activity in inhibiting release of calcium from bone and in
				treating e.g. osteoporosis, Paget's disease, bone metastases, or
				destructive rheumatoid arthritis.
				Claim 5: Page 75-77: 96pp: English.
				DR PN WO988423-A2.
				XX
				PD 02-JUL-1998.
				XX
				PF 18-DEC-1997;
				XX
				PR 20-DEC-1996;
				XX
				PR 96US-0772441.
				XX
				(TEXA) UNIV TEXAS SYSTEM.
				XX
				PI Choi S, Reddy SV, Goodman GD;
				XX
				WPI: 1998-377654/32.
				DR N-PSDB: ANV44866.
				XX
				This sequence represents the human osteoclast inhibitor protein, OIP-2
				which can be used in the method of the invention. The method is
				for inhibiting osteoclastogenesis, and comprises administering to a ce
				an osteoclast (OC) inhibitor protein (OIP) composition in a vehicle
				where the composition inhibits production of OCs. The OIP polypeptides
				have activity in inhibiting release of calcium from bone and in
				treating e.g. osteoporosis, Paget's disease, bone metastases, or
				destructive rheumatoid arthritis.

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 27, 2003, 15:21:29 ; Search time 38 Seconds (without alignments)

Title: US-09-934-066-2

Perfect score: 2454

Sequence: 1 MSSPLGHFQILVFLHALLIF.....CNMGVDVKQTIVSAIEACSM 466

Scoring table: BL05UM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0*

Maximum Match 100*

Listing first 45 summaries

Database : SPTRMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rabbit:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archeap:*

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2454	100.0	466	10	Q9LJX8		Q9LJX8 arabidopsis
2	1266.5	51.6	483	10	Q9XFZ4		Q9XFZ4 vigna mungo
3	1254.5	51.1	483	10	Q9AUD9		Q9AUD9 phaseolus a
4	1252	51.0	482	10	Q9XFZ5		Q9XFZ5 vigna mungo
5	1246	50.8	494	10	Q93VM0		Q93VM0 arabidopsis
6	1243.5	50.7	484	10	Q243Z5		Q243Z5 phaseolus v
7	1236	50.4	486	10	Q949L7		Q949L7 beta vulgaris
8	1228.5	50.1	501	10	Q9SSZ4		Q9SSZ4 oryza sativa
9	1219	49.7	492	10	Q9M4R6		Q9M4R6 ipomoea batatas
10	1208	49.2	485	10	Q9ZPZ8		Q9ZPZ8 zea mays (m)
11	1204.5	49.1	485	10	Q9FER6		Q9FER6 zea mays (m)
12	1191.5	48.6	481	10	Q9SBX2		Q9SBX2 zea mays (m)
13	1183.5	48.2	486	10	Q9SBX3		Q9SBX3 zea mays (m)
14	1182.5	48.2	486	10	Q9FER7		Q9FER7 zea mays (m)
15	1166.5	47.5	460	10	Q9SM0		Q9SM0 lycopersico
16	1160	47.3	486	10	Q93VS7		Q93VS7 arabidopsis

ALIGNMENTS

RESULT 1	ID	Q9LJX8	PRELIMINARY,	PRT,	466 AA.
	ID	Q9LJX8;			
	AC	Q9LJX8;			
	DT	01-OCT-2000 (TREMBLrel. 15, Created)			
	DT	01-JUN-2002 (TREMBLrel. 15, Last sequence update)			
	DE	vacuolar processing enzyme (proteinase).			
	OS	Arabidopsis thaliana (Mouse ear cress).			
	OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
	OC	NCBI_TaxID=3702;			
	RN	[1]			
	RP	SEQUENCE FROM N.A.			
	RC	STRAIN=COLUMBIA;			
	RA	Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.; Submitted (JUN-1999) to the EMBL/CenBank/DDAJ databases.			
	RL	[2]			
	RN	SEQUENCE FROM N.A.			
	RP	SEQUENCE FROM N.A.			
	RC	MEDLINE=20363399; PubMed=10907853;			
	RA	Nakamura Y.;			
	RT	"Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the regions of 4,251,695 bp covered by ninety P1, TAC and BAC clones."			
	RT	TAC and BAC clones.			
	RL	DNA Res. 7:217-221(2000).			
	DR	EMBL: AP000383; BAB01880.1; InterPro: IPR001096; Legumain; Peptidase.C13; 1.			
	DR	PF01650; HEMOGLOBINASE; PR00776; Peptidase.C13; 1.			
	SQ	SEQUENCE 466 AA; 52055 MW;			

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	2454	100.0	466	10	Q9LJX8		Q9LJX8 arabidopsis
2	1266.5	51.6	483	10	Q9XFZ4		Q9XFZ4 vigna mungo
3	1254.5	51.1	483	10	Q9AUD9		Q9AUD9 phaseolus a
4	1252	51.0	482	10	Q9XFZ5		Q9XFZ5 vigna mungo
5	1246	50.8	494	10	Q93VM0		Q93VM0 arabidopsis
6	1243.5	50.7	484	10	Q243Z5		Q243Z5 phaseolus v
7	1236	50.4	486	10	Q949L7		Q949L7 beta vulgaris
8	1228.5	50.1	501	10	Q9SSZ4		Q9SSZ4 oryza sativa
9	1219	49.7	492	10	Q9M4R6		Q9M4R6 ipomoea batatas
10	1208	49.2	485	10	Q9ZPZ8		Q9ZPZ8 zea mays (m)
11	1204.5	49.1	485	10	Q9FER6		Q9FER6 zea mays (m)
12	1191.5	48.6	481	10	Q9SBX2		Q9SBX2 zea mays (m)
13	1183.5	48.2	486	10	Q9SBX3		Q9SBX3 zea mays (m)
14	1182.5	48.2	486	10	Q9FER7		Q9FER7 zea mays (m)
15	1166.5	47.5	460	10	Q9SM0		Q9SM0 lycopersico
16	1160	47.3	486	10	Q93VS7		Q93VS7 arabidopsis

Query Match	Best Local Similarity	Length 466;
Matches 466;	Conservative 0;	Mismatches 0;
		Indels 0;
		Gaps 0;
QY	1 MSSPLGHFQILVFLHALLFSAESRKTOILNDVESSKSAGCTRWAVLGSSNEYNN	60

Db	1	MSSPLGHQFQILFLHALLIFLTSRSRSTQOLLNDVSSSDISAGTRWAVLAGSNEYNN	60	Qy	124	AVNQNFYNVLGNESGYTGGNGKVKSPNDNIFTYADHGAPGLIAMPITGDEYMAKF	183
Db	61	RHOADICHAYQILRKGGKLDENIIVPMDDIAFSENPRGVINKPDKDGYKGVPKDY	120	Qy	126	DYIVNNFAAFLGNSKAISLTTGCGSMFEGIUKKNUINTYAVTAANSKESSWGTCPDN	185
Db	61	RHOADICHAYQILRKGGKLDENIIVPMDDIAFSENPRGVINKPDKDGYKGVPKDY	120	Qy	184	NEVLEKMKRKKYKVNKVITYVEACSGSMFEGIUKKNUINTYAVTAANSKESSWGTCPDN	243
Db	121	TKIAVNVQFYAVNLGNESGYTGGNGKVKSPNDNIFTYADHGAPGLIAMPITGDEVA	180	Qy	186	LEVKKRKAASGTYKSLAFLEGESGTFGGULPEGINITYATTAANKESSWGTCPDN	245
Db	121	TKEAVNQFYAVNLGNESGYTGGNGKVKSPNDNIFTYADHGAPGLIAMPITGDEVA	180	Qy	244	PPPSECTCLGDTFTSMLEOSDLHDMSKETLEQQHYKVR - VGSVDVPTSHVCREGT	302
Db	181	KDFNEVLEKMKRKKYKVNKVITYVEACSGSMFEGIUKKNUINTYAVTAANSKESSWGTCP	240	Qy	246	PSPPPEYETCLGDSYAWMEDSDIHNRTELIQOPELVAORTMNGNSAVGSHMVOYGD	305
Db	181	KDFNEVLEKMKRKKYKVNKVITYVEACSGSMFEGIUKKNUINTYAVTAANSKESSWGTCP	240	Qy	303	EKMFDKLYSLSYIGRNPNENDNFETESPSSPIISNGSLVNPDIPLYLQRKIQKAPMGSL	362
Qy	241	ESYPPPSEIGTCLGDTFSISWLEDSDLHDMSKETLEQQHYKVRKERVSDVPTSHVCRF	300	Qy	306	VGLSKNNVSYLQTNPNDNEPREK - NSLVPSKAVNQRDAVLVWDFDKPKAPLGSSR	364
Db	241	ESYPPPSEIGTCLGDTFSISWLEDSDLHDMSKETLEQQHYKVRKERVSDVPTSHVCRF	300	Qy	363	SKEAQKQLLDEKHNRKQDQSITDILRSVKTNVNLLTSTRTGQPLVDDWDCFKTLV	422
Qy	301	GTEKMLADYSLSYIGRNPNENDNFETESPSSPIISNGSLVNPDIPLYLQRKIQKAPMG	360	Qy	365	KSYAQKQIILEAMSHRMHIDSVTFLIGKLFEGPULLSSYVRPAGQPLVDDWDCFKTLV	424
Db	301	GTEKMLADYSLSYIGRNPNENDNFETESPSSPIISNGSLVNPDIPLYLQRKIQKAPMG	360	Qy	423	NSFRNHGCGATHVHGLKYTGALANICCNNGDVYKQTVSAIEQAC	464
Qy	361	LESKEAQKQLLDEKHNRKQDQSITDILRSVKTNVNLLTSTRTGQPLVDDWDCFKT	420	Db	425	RTEFHCGQLSQQGMKHMRSFANLCNAGIRKQMAEASAQAC	466
Db	361	LESKEAQKQLLDEKHNRKQDQSITDILRSVKTNVNLLTSTRTGQPLVDDWDCFKT	420				
				RESULT	3		
				Q9AID9			
				ID	O9AID9	PRELIMINARY:	
				AC	O9AID9;	PRT:	483 AA.
				DT	01-JUN-2001 (TREMBLrel.		
				DT	01-JUN-2001 (TREMBLrel.	17.	Created)
				DT	01-MAR-2002 (TREMBLrel.	17.	Last sequence update)
				DT	01-MAR-2002 (TREMBLrel.	20.	Last annotation update)
				DE	Asparaginyl endopeptidase.		
				GN	SEQUENCE FROM N.A.		
				RN	TAINAN 5; TISSUE=COTYLEDON;		
				RP	Phaeolus aureus (Mung bean) (Vigna radiata).		
				RC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
				OC	Strang W.H., Chen C.S.;		
				OC	Spermatophyta; Magnoliophyta; eudicots; Rosidae;		
				OC	euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Vigna.		
				NCBI_TAXID	NCBI_TAXID=3916;		
				RN	[1].		
				DR	DR MEROPS: C13.002;		
				DR	DR InterPro: IP001006; Legumain.		
				DR	DR PFam: PF01650; Peptidase_C13; 1.		
				DR	DR PRINTS: PR0076; HEMOGLOBINASE.		
				SQ	SEQUENCE 483 AA; 53305 MW; 733B7E283EF9B9264 CRC64;		
					Query Match 51.1%: Score 1254; DB 10; Length 483;		
					Best Local Similarity 54.1%: Prcd. No. 3, 8e-81;		
					Matches 231; Conservative 71; Mismatches 124; Indels 1; Gaps 1;		
					Qy 38 SDSAKGTTRAVLIVGNSNEYNYQYQDICHAYQTLRKGGJLDRNTIVMFYDDIAFSSEN	97	
					97		
					Db 41 NDDNWKTRWILFAGSNGWNYWRHQDICHAYQTLRKGGJLKEENIVYDDIAFNWN	100	
					100		
					Query Match 51.1%: Score 1254; DB 10; Length 483;		
					Best Local Similarity 54.1%: Prcd. No. 3, 8e-81;		
					Matches 231; Conservative 71; Mismatches 124; Indels 1; Gaps 1;		
					Qy 98 PRGGVINKPGDGDYKGPVDPYTKAEVNVORFYNVLGNESGVGGNGKVKSPNDN	157	
					157		
					Db 101 PRGGVINKPGDGDYKGPVDPYTKAEVNVORFYNVLGNESGVGGNGKVKSPNDN	160	
					160		
					Qy 158 FLYADICAPGLIAMPIDEVAKDFNEVLEKMIKRKKYKNMVIVVEACSGSMPEGILK	217	
					217		
					Db 161 FLYSDRGGPGVGTGPGYTPAGVYIASLVEVLRKKHASCTYKLVYIACEAGSIFGCLIP	220	
					220		
					Qy 218 KNINIIYAVTAANSKESSWVGVCPCESYPPPSPBEGTCLGDTFSISLEDSDLHDMSKETLE	277	
					277		
					Db 221 EDNINIIYATASNEESESSYGTGEPYSPSYCCTGEGYPSYPCSTGEGYPSYPCSTGEGYPSYPC	280	
					280		
					Qy 278 QOYHVVKRGGSDYPTSHVCREGTERMMLKDYLSSTYGRNPNDNETTESFSSPLNSNC	337	
					337		

Db	281	QQYKVKDRTLSGGWYQGSHVMOYGVDFESKDTLFLYLGTDANDNLTFVDE-NSLMSSST	339	Qy	367	QKKLLEKKNHRKQIDOSITDILRSVKQTNVNLNTSTRTGQPLVYDDWDCFKTLVNSFK 426
Qy	338	LNPNDPFLYLQRKIKAPGSLESKEAQKKLDEKKNHRKQIDOSITDILRSVKQTNV	397	Db	368	RKVLEVMSRMHIDDSVKLKGKLLFGFEKAPEVLNAVRPGSALVDDWACLKTMVRTE 427
Db	340	AVNQRDADLVFWHKRKAPEPKPNKEARQVLEVMSHRMHIDDSVKLGFGEKA	399	Qy	427	NHCATVHYGLKYTGALANICNMGVDFVKOTVSIAEQAC 464
Qy	398	LNLITSTRTRCQLPVDWDCEKTLVNSFKHICATHYGLKYTGALANICNMGVDFVKQT	457	Db	428	THCGSLSQGMKHMSPFANICNGVKEQMAEASAQC 465
Db	400	PEVNAVRPAGSAVLVDWACLKTMVRTEHGSSLSQYGMKHMRSFANICNGVKEQMA	459			
Qy	458	SAIEQAC 464				
Db	460	EASQAQC 466				
			RESULT 5			
			Q93VM0			
			ID Q93VM0	PRELIMINARY;	PRT;	494 AA.
			AC Q93VM0;			
			DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
			DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
			DE Putative gamma-VE protein (At4G32940/F26P21_60).			
			GN AT4G32940			
			OS Arabidopsis thaliana (Mouse-ear cress).			
			OC Eukaryota; Viridiplantae; Streptophytta; Embryophyta; Tracheophyta;			
			OC Spermatophyta; Magnoliophyta; eudicots; Rosidae;			
			OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.			
			OX NCBI_TAXID=3702;			
			RN [1]			
			RP SEQUENCE FROM N.A.			
			RA Yamada K., Banh J., Banno F., Chang E., Dale J.M., Goldsmith A.D.,			
			RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,			
			RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,			
			RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,			
			RA Kawai J., Kim C., Koeseema E., Lam B., Lin J., Meyers M.C., Miranda M.,			
			RA Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,			
			RA Shin P., Southwick A., Tracy S.E., Shinozaki K.,			
			RA Ecker J.R., Theologis A.;			
			RA Full Length cDNA of gene AT4G32940 (GI:7270241).;			
			RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
			RN [2]			
			RP SEQUENCE FROM N.A.			
			RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,			
			RA Goldsmith A.D., Lee J.M., Quach H.L., Tang C., Toriumi M., Yu G.,			
			RA Bowser L., Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J.,			
			RA Jones T., Kamiya A., Karlin-Neumann G., Kawai J., Kim C.,			
			RA Lam B., Lin J., Meyers M.C., Miranda M., Narusaka M., Nguyen M.,			
			RA Palm C.J., Sakurai T., Satou M., Seki M., Shin P., Southwick A.,			
			RA Shinozaki K., Davis R.W., Ecker J.R., Theologis A.;			
			RT Full Length cDNA of gene AT4G32940 (GI:7270241).;			
			RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.			
			RN [3]			
			RP SEQUENCE FROM N.A.			
			RA Shin P., Chen H., Cheuk R., Kim C.J., Koeseema E., Meyers M.C.,			
			RA Goldsmith A.D., Lee J.M., Carninci P., Dale J.M., Goldsmith A.D.,			
			DR MEROPS; C13_002; -.			
			DR InterPro; IPR001096; Legumain.			
			DR PRINTS; PR00776; HEMOGLOBINASE.			
			SQ SEQUENCE: 482 AA; 53211 MW; C8262E868ED378E8 CRC64;			
			Query Match Similarity 51.0%; Score 1252; DB 10; Length 482;			
			Best Local Similarity 51.5%; Pred. No. 5.7e-81; Indels 8; Gaps 3;			
			Matches 236; Conservative 78; Mismatches 136; Insertions 136;			
			DR LYLHALLIFSSERKRT---QLNDNVESSDKSAGKTRAVLVAWSNEYNNRHQADI 66			
			Qy 11 LYLHALLIFSSERKRT---QLNDNVESSDKSAGKTRAVLVAWSNEYNNRHQADI 66			
			DR 12 LLFLATVALVAARDHVGDFLRLPSDS--GNDVQGTRPAILEFGNSGYNNRHQADI 68			
			Qy 67 CHAYQIYLKGGLKIDENTIVMDIAFSSENPRGVTINKPDGEDYKGVPKDYGTEKAVN 126			
			DR 13 CHAYQIYLKGGLKIDENTIVMDIAFSSENPRGVTINKPDGEDYKGVPKDYGTEKAVN 126			
			DR 69 CHAYQIYLKGGLKIDENTIVMDIAFSSENPRGVTINKPDGDYEGVPKDYGTEDAT 128			
			Qy 127 VQNFYNVLGNESGTVGGNGKVKSGPNDNFIYIYADHGAPGLIAMPTGEVMAKDFNEV 186			
			DR 129 AHNFYSALLGDSALJGGSGKVSSGDDRFYFSDHGGPGVLPAGPYIASLVEV 188			
			Qy 187 LEKMHKRKKYKMKVIYVEACESGSMSEGILKKNNIYATAKSSESSWGVYCPESYPPP 246			
			DR 189 LKKHAGSGTYKLNFVLEACEAGSTEGLIPEDINYATPASNAAESSWGVYCPESYPPP 248			
			Qy 247 PSEIGCLGDFISIWEDSLHDMSKETEQHVVKRVGSDYETSHCVRCFTCKML 306			
			DR 249 PEYSYCLGDLYSVAMMEDSRHNLRTESLHQKYKVKDRTLSGGHYSVQGDVEFS 308			
			Qy 307 KDYLSSYGRNPNDNFTETSFSSPISNSGLIVPRDIPLYLQRRIKQAPMGLESKEA 366			
			DR 309 KDAFLYLGTDANDLTFTVDE-NSLWSSSTAVNQRDADLVHFWHKRKAEGSPRKNEA 367			
			Query Match Score 50.8%; Best Local Similarity 53.6%; DR 10; Length 494;			
			Matches 233; Conservative 71; Mismatches 121; Indels 10; Gaps 4;			
			Qy 36 ESSDKSAKGTRAVLYAGSNENYYRHODICHAYQTLRKGLKDENIVMFYDDIAFS 95			
			DR 49 ENDDDSNSGTRAVLYAGSGWNYRHODICHAYQTLRKGLKDENIVMFYDDIANY 108			

GS	SEE2A. Zea mays (Maize)	OC	Asteridae; euasterids I; Solanales; Solanaceae; Solanum
OS	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	OX	NCBI_TaxID=4081;
OC	Panicoidea; Andropogoneae; Zea.	RN	SEQUENCE FROM N.A.
OC	"Characterisation of See2, a senescence enhanced cDNA from maize with homology to legumin."	RA	STRAIN-CV. WEST VIRGINIA 106; TISSUE=FRUIT;
OX	Submitted (DEC-199) to the EMBL/GenBank/DBJ databases.	RC	Lemaire-Chamley M., Petit J., Raymond P., Chevalier C.; "Isolation and characterisation of a fruit specific cDNA clone for vacuolar processing enzyme from tomato (accession No. AJ243876).
RN [1]	EMBL: A243876; CAB51345; Peptidase_C13_1.	RT	RT (PGH9-164).";
RP	EMBL: IPB001096; Legumin_Pfam1; Peptidase_C13_1.	RL	RT (121:1057-1057 (1999)).
RA	InterPro: IP001096; Legumin_Pfam1; Peptidase_C13_1.	DR	EMBL: A243876; CAB51345..1..;
RA	InterPro: IPB001096; Legumin_Pfam1; Peptidase_C13_1.	DR	InterPro: IPB001096; Legumin_Pfam1; Peptidase_C13_1.
RT	PRINTS: PR00776; HEMOGLOBINASE.	DR	PRINTS: PR00776; HEMOGLOBINASE.
RT	SEQUENCE FROM N.A.	DR	SEQUENCE 460 AA; 51801 MW; 469EC61A01FFC6DD CRC64;
RA	STRAIN-CV. MISSOURI 17 INBRED LINE; Donnison I.S., Griffiths C.M., Thomas A., Hosken S.E., Bridges I., Thomas H.; "Characterisation of See2, a senescence enhanced cDNA from maize with homology to legumin."	DR	Score 1166.5; Pred. No. 6.5e-75;
RT	Submitted (DEC-199) to the EMBL/GenBank/DBJ databases.	DR	Best Local Similarity 51.1%; Indels 17; Gaps 5;
RT	EMBL: A251453; CAC18099..1..;	SO	Query Match 47.5%; Score 1166.5; Pred. No. 6.5e-75;
DR	MEROPS; C13..002; -	Qy	13 FJ1ALLIFSAESRKTQLINDVESSRSKSACTRIVAVLVAQSNEYYNRIOADICHAYQI
DR	InterPro: IPB001096; Legumin_Pfam1; Peptidase_C13_1.	Db	10 FJ1ALLIFSAESRKTQLINDVESSRSKSACTRIVAVLVAQSNEYYNRIOADICHAYQI 65
DR	PRINTS: PR00776; HEMOGLOBINASE.	Qy	14 LRKGGLKDEHLLIVFMDIANPENPRPGVLTNNPPIGHDVKGVPDVTGHDCAONFY 132
KW	SIGNAL.	Db	66 LRKGGLKDEHLLIVFMDIANPENPRPGVLTNNPPIGHDVKGVPDVTGHDCAONFY 125
FT	1 20 SIGNAL.	Qy	15 VLGKNECYTGCGNGKVKSGPNDNFVLYADUGAPGLIAMPGDEMAKDNEVLEKMKH 192
CHAIN	21 486 AA; 53428 MW; 19C10A341DD33BF5 CRC64;	Db	126 VLGKNSAIIIGGAKVUNSGPNDYIIFTYTYDAGPLVGMEDPPYAIIDNEVKKHA 185
SQ	SEQUENCE 486 AA; 53428 MW; 19C10A341DD33BF5 CRC64;	Qy	193 RKYKNMVITYVEACESGSMFEGITLKKNNIUYAVTAANSKESSWGVYC-----PESYP--P 245
Query Match 48.2%; Score 1182.5%; DB 10; Length 486;	Db	186 SRYKKMVFYEAQDSSMFADLDEGENIATTSKSKPDEGOWATCYFPTDTSCYFGCBP 245	
Best Local Similarity 51.8%; Pred. No. 5.1e-76;	Qy	194 PPSEIGTCIGDFTFSIWLEDSDLHDMSKETLEQOYHVVKRVRGSDYPTSHVCRFCTEKM 305	
Matches 25%; Conservative 68; Mismatches 134; Indels 7; Gaps 3;	Db	246 PKDFKDNCLGDLPSVSNLENSDFLHDLQVETJEKOYLRIHKRVLNNGTGSIMYQVGDLLH 305	
Qy	36 FSSDKSAKGCTPRAVLYAGSNEYYNYRHOADICHAQOILRKGGLKDENITIYMDIAFSS 95	Qy	306 LKDYLSSYIGRNPENDNFTTSFSSPITSNSGLVNPRDIPLYLORKTIQAPMGSLESKE 365
Db	40 DEFDGDVGTRAVLIAQSSGYVNYRHOADICHAQOIMKGGKLQDENITIYMDIAHSP 99	Db	306 NRDALSTIMGNSNPKHTWS -- ANNNNASNSHVNORDVQVLKLIKFRQNPNEGSRRKNE 362
Qy	96 PRPGVTTINKPGDEDYKGPVDYKTDYKEAVNVQONYNVILGNESCVGGNGKVKSGPND 155	Qy	366 AOKKLDEKHNKRQIDOSITDILRLSKVOTVNLJLUTSTRTGOPLEVVDDNCFCFTLVNSF 425
Db	100 ENRPGVTTNHQQGGDYAGVPKDTGROYVNNDNEFAVLLGNKTAIRGGSKVVDGPND 159	Db	363 AYRLLSEVSEREHVDKSVKHGQILFGVENGQKVNTVR -- QPLVDDNDCLKSEFVKF 419
Qy	156 NITIYADHGAGPLIAMPGDEVIMARDFNEVLERMHKRKTKYKNTVYEACESGSMFEGI 215	Qy	426 KNHGATHYGLRTGALANICNMGDVQTVSAIEOACS 465
Db	160 HIVFVSDHGGQVGLGRPTYLYGDDLVLDVLYKHAAGTYKSYVYLEACESSSIIPEGI 219	Db	420 ESHCGSLTSYGGKHHVRFGANMCNAGIORDQMDAAKOTCS 459
Qy	216 LKKNLNIYAVTAANSKESSWGVYCPESYPPPEIGTCGDTFSISWLEDSDLHDMSKET 275	Search completed: May 27, 2003, 15:27:37	
Db	220 LPDIINVATTNAEADSESWGTYCPGEFPSSPPPEIGTCGDTFSISWLEDSDLHDMSKET 279	Job time : 41 secs	
Qy	276 LEOQHYVVKRVSQDSVP --- SHYCRFSTPEKMRKDYLSSYIGRNPDENDNFTTESFSS 331.		
Db	280 LKQKQYLVKDRPA -> HDTFSYGSYHMQYGALELNVORESYIGTPANDGNTFEDNSL 337		
Qy	332 PTNSGLVNPDRFLYLRQTKQAPMGSLSESKEAGQKLDEKHNKRQIDOSITDILRLS 391.		
Db	338 P-SFSKAVNQRDADLYFWQKRYKLADSSHAKNEARKEELVMARSHDVSSEVIGSSL 396		
Qy	392 VKGTNVNLNLITSTRTPGOPLWDGCFKTLVNSFKHCGATVHGLKYGKALANICNMGV 451.		
Db	397 FGSEDGPRVLRKVARAPEPLVDDWSCLSKTSVRTFEARCGLSALQTCMKHRSFANNCNACT 456		
Qy	452 DRYQTVSAEQNCS 465		
Db	457 LPPAVSKVAAQACCS 470		
RESULT 15			
Q9SMDO	PRELIMINARY;	PRT:	460 AA.
ID			
Q9SMDO;			
AC			
DT	01-MAY-2000 (TREMBrel. 13, Created)		
DT	01-MAY-2000 (TREMBrel. 13, Last sequence update)		
DT	01-JUN-2002 (TREMBrel. 21, Last annotation update)		
DE	Vacuolar Processing enzyme.		
GN			
OS	Lycopericon esculentum (Tomato).		
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;		

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model.

Run on: May 27, 2003, 15:25:48 ; Search time 17 Seconds
 (without alignments)

Title: US-09-934-066-2
 Perfect score: 2444
 Sequence: 1 MSSPLGHFQILVFLHALLIF.....CNMGVDVKQTIVSATAEQACSM 466

Scoring table: BLOSUM62
 Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
 Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_AA:*

1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*

2: /cn2_6/ptodata/1/iaa/5B_COMB.pep:*

3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*

4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*

5: /cn2_6/ptodata/1/iaa/PCUTS_COMB.pep:*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	828	33.7	433	2	US-09-139-424-4
2	828	33.7	433	4	US-08-706-216-4
3	744	30.3	431	2	US-08-928-613-2
4	744	30.3	431	3	US-09-193-524-2
5	137	5.6	619	4	US-09-134-001C-4248
6	115.5	4.7	3169	4	US-09-453-702B-57
7	114.5	4.7	580	4	US-09-367-583-2
8	112.5	4.6	1440	4	US-09-357-251-37
9	112.5	4.6	1512	4	US-09-443-184-48
10	103	4.2	1338	1	US-08-471-033-50
11	103	4.2	1338	2	US-08-471-044-50
12	103	4.2	1338	2	US-08-463-483A-50
13	103	4.2	1338	2	US-08-471-046A-50
14	103	4.2	1338	2	US-08-470-566B-50
15	103	4.2	1338	3	US-08-469-534-50
16	103	4.2	1338	3	US-09-300-529-50
17	100	4.1	877	4	US-09-619-353-2
18	97	4.0	503	1	US-08-245-294-8
19	97	4.0	503	1	US-08-474-499-8
20	97	4.0	503	1	US-08-307-279A-8
21	97	4.0	503	4	US-09-555-310-8
22	97	4.0	503	5	PCT-US95-06211-8
23	96	3.9	906	4	US-09-134-001C-3218
24	96	3.9	15281	2	US-08-471-119A-2
25	94.5	3.9	1479	4	US-08-951-912-4
26	94.5	3.9	1479	4	US-09-174-077-4
27	94.5	3.9	3135	1	US-08-323-170B-2

ALIGNMENTS

RESULT 1
 US-09-139-424-4
 ; Sequence 4, Application US/09139424
 ; Patent No. 5985832
 ; GENERAL INFORMATION:
 ; APPLICANT: Rodman, G. David
 ; APPLICANT: Reddy, Sakamuri V.
 ; APPLICANT: Choi, Sun-Jae
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE FOR
 ; OSTEOCLAST INHIBITORY FACTOR
 ; NUMBER OF SEQUENCES: 4
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Arnold, White & Durkee
 ; STREET: P.O. Box 4433
 ; CITY: Houston
 ; STATE: Texas
 ; COUNTRY: USA
 ; ZIP: 77210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/139-424
 ; FILING DATE: 2003-09-13
 ; CLASSIFICATION:
 ; PRIORITY APPLICATION DATA:
 ; APPLICATION NUMBER: 08/772,441
 ; FILING DATE: 2001-07-12
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Kitchell, Barbara S.
 ; REGISTRATION NUMBER: 33,928
 ; REFERENCE/DOCKET NUMBER: UTSK-295
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 512/418-3000
 ; TELEFAX: 512/474-7577
 ; INFORMATION FOR SEQ ID NO: 4:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 433 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-139-424-4

Result No.	Score	Query Match	Length	DB ID	Description
1	828	33.7	433	2	Sequence 4, Appli
2	828	33.7	433	4	Sequence 4, Appli
3	744	30.3	431	2	Sequence 2, Appli
4	744	30.3	431	3	Sequence 2, Appli
5	137	5.6	619	4	Sequence 4248, Appli
6	115.5	4.7	3169	4	Sequence 257, Appli
7	114.5	4.7	580	4	Sequence 2, Appli
8	112.5	4.6	1440	4	Sequence 37, Appli
9	112.5	4.6	1512	4	Sequence 48, Appli
10	103	4.2	1338	1	Sequence 50, Appli
11	103	4.2	1338	2	Sequence 50, Appli
12	103	4.2	1338	2	Sequence 50, Appli
13	103	4.2	1338	2	Sequence 50, Appli
14	103	4.2	1338	2	Sequence 50, Appli
15	103	4.2	1338	3	Sequence 50, Appli
16	103	4.2	1338	3	Sequence 50, Appli
17	100	4.1	877	4	Sequence 2, Appli
18	97	4.0	503	1	Sequence 8, Appli
19	97	4.0	503	1	Sequence 8, Appli
20	97	4.0	503	1	Sequence 8, Appli
21	97	4.0	503	4	Sequence 8, Appli
22	97	4.0	503	5	Sequence 8, Appli
23	96	3.9	906	4	Sequence 3218, Appli
24	96	3.9	15281	2	Sequence 2, Appli
25	94.5	3.9	1479	4	Sequence 4, Appli
26	94.5	3.9	1479	4	Sequence 4, Appli
27	94.5	3.9	3135	1	Sequence 2, Appli

Query Match 33.78; Score 828; DB 2; Length 433;
 Best Local Similarity 38.28; Pred. No. 7.5e-73;
 Matches 179; Conservative 80; Missmatches 157; Indels 52; Gaps 13;
 Qy 8 FQILVFHLALLIFSAERSRKTOLLNDYVESSDKSAKGTRAVLVAGSNYYNNRHOADC 67

Query Match Score: 828; DB: 4; Length: 433;
 Best Local Similarity: 38.2%; Pred. No.: 7.5e-73;
 Matches: 179; Conservative: 80; Mismatches: 157; Indels: 52; Gaps: 13;

Qy 8 FQILVPHALLIFSAERSRKTOOLLNDVESSDKSAGTRAVLVAWSNEYNNYRHADIC 67
 Db 3 WKAFLVALGIGA-----VPIDDPEDGKHHWIVACSGNYNRYIQAAC 50
 Qy 68 HAYOILRKGGGLDENITIVMDIAFSENPRPGVFLINKPDGEDYKGVPKDYTEKEAVNV 127
 Db 51 HAYOILHNRGIPDEQIVMMYDIASEDNPYPGIVINRPNTDVGVPKDYGEDVTP 110
 Qy 128 QNFPIVNLGNECSVTG-CNGKVVKSGPNDNIFTIYADHGAPFLIAMPTGDEVIAKDFNEV 186
 Db 111 ONPLAVLRGDAEAVKGIGSKVLSKGPDSVHFTYDHFSTGLVVF-NEDLHYKDLNET 169
 Qy 187 LEMMHKKRKKYKNNMVIYEAECGSMPEGILKKNLNLYAVTAANSKESSMGVYCPESYPPP 246
 Db 170 IHWYKHKYMRKMFYEAEGCSMMH-LFDNINVYATTAAANPRESSTACYDE-----223
 Qy 247 PSEIGTCGLDTFSISLMSLSDLSLHDMSKETLBOQHVVKRGSDVPESTSHVCRFGTEKML 306
 Db 224 -KRSTLGDWNNSVNMFEDSDVLDLKLTHQYLHVKSHT----NTSHVMOGNKTIS 276
 Qy 307 KDLSSYTGPNPNDNFTESFSS----PISSNCLVNPNRDPDILYLQRKIQKAPMGSIL 361
 Db 277 TMKVMQFG-----MKRKASSPVPLPPVTHLDLTPSPDVPLTMRKLMNT--NDL 325
 Qy 362 -ESKEAQKKLLEDEKHARKQDOSITDLRLSVKQTVNLNLTSRPTGQPLVDDMDCFKT 420
 Db 326 EBSRQLTTEEIQHRLIEKSVRKVSLIAASEAEQELLSERA--PLTGH-SCYPE 381
 Qy 421 LVNSFKHIC----GATVHYGKTYGALANICNMGVDYKQTVSATEONC 464
 Db 382 ALLHFRTHCFNWHSPTEYALRHLVYLVNCQEKPYPLHRIKLSDHVC 429

RESULT 2
 US-08-706-216-4
 Sequence 4 Application US/08/06216
 GENERAL INFORMATION:
 PATENT NO.: 610098
 NUMBER OF SEQUENCES: 6
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: DNX Research Institute
 STREET: 901 California Avenue
 CITY: Palo Alto
 STATE: California
 COUNTRY: USA
 ZIP: 94304-1104
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PATENT RELEASE #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/706,216
 FILING DATE: 30-AUG-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Chang, Edwin P.
 REGISTRATION NUMBER: 34,090
 REFERENCE/DOCKET NUMBER: DX0613
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-852-9196
 TELEFAX: 415-96-1200
 INFORMATION FOR SEQ ID NO: 4:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 433 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein

Query Match Score: 33.7%; DB: 4; Length: 433;
 Best Local Similarity: 38.2%; Pred. No.: 7.5e-73;
 Matches: 179; Conservative: 80; Mismatches: 157; Indels: 52; Gaps: 13;

Qy 68 HAYOILRKGGGLDENITIVMDIAFSENPRPGVFLINKPDGEDYKGVPKDYTEKEAVNV 127
 Db 51 HAYOILHNRGIPDEQIVMMYDIASEDNPYPGIVINRPNTDVGVPKDYGEDVTP 110
 Qy 128 QNFPIVNLGNECSVTG-CNGKVVKSGPNDNIFTIYADHGAPFLIAMPTGDEVIAKDFNEV 186
 Db 111 ONPLAVLRGDAEAVKGIGSKVLSKGPDSVHFTYDHFSTGLVVF-NEDLHYKDLNET 169
 Qy 187 LEMMHKKRKKYKNNMVIYEAECGSMPEGILKKNLNLYAVTAANSKESSMGVYCPESYPPP 246
 Db 170 IHWYKHKYMRKMFYEAEGCSMMH-LFDNINVYATTAAANPRESSTACYDE-----223
 Qy 247 PSEIGTCGLDTFSISLMSLSDLSLHDMSKETLBOQHVVKRGSDVPESTSHVCRFGTEKML 306
 Db 224 -KRSTLGDWNNSVNMFEDSDVLDLKLTHQYLHVKSHT----NTSHVMOGNKTIS 276
 Qy 307 KDLSSYTGPNPNDNFTESFSS----PISSNCLVNPNRDPDILYLQRKIQKAPMGSIL 361
 Db 277 TMKVMQFG-----MKRKASSPVPLPPVTHLDLTPSPDVPLTMRKLMNT--NDL 325
 Qy 362 -ESKEAQKKLLEDEKHARKQDOSITDLRLSVKQTVNLNLTSRPTGQPLVDDMDCFKT 420
 Db 326 EBSRQLTTEEIQHRLIEKSVRKVSLIAASEAEQELLSERA--PLTGH-SCYPE 381
 RESULT 3
 US-08-908-613-2
 Sequence 2 Application US/08/28613
 GENERAL INFORMATION:
 PATENT NO.: 580562
 APPLICANT: Diep, Dinh
 APPLICANT: Braxton, Scott M.
 APPLICANT: Deleane, Angelo M.
 TITLE OF INVENTION: A NOVEL HUMAN CYSTEINE PROTEASE
 CORRESPONDENCE ADDRESS:
 ADDRESS: INCYTE PHARMACEUTICALS, INC.
 STREET: 3174 Porter Drive
 CITY: Palo Alto
 STATE: CA
 COUNTRY: USA
 ZIP: 94304
 COMPUTER READABLE FORM:
 MEDIUM TYPE: FLOPPY DISK
 COMPUTER: IBM PC COMPATIBLE
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PATENT RELEASE #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/928,613
 FILING DATE: 11-JUN-1996
 CLASSIFICATION: 524
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/567,506
 ATTORNEY/AGENT INFORMATION:
 NAME: Luther, Barbara J.
 REGISTRATION NUMBER: 33354
 REFERENCE/DOCKET NUMBER: PF-0048 US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-855-0555

; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-928-613-2

Query Match 30.3%; Score 744; DB 2; Length 431;
Best Local Similarity 38.4%; Pred. No. 1.4e-64;
Matches 166; Conservative 70; Mismatches 128; Indels 68; Gaps 14;

Qy 8 FOLYLHALIFSAESRKTOQLLNDVSESSDKSAGTRWAVLAGSNEYNNYRHODIC 67
Db 3 WRVVVFSLVAIGA ----- VPIDPEDGKHHWVIVAGSGNWNYRQADAC 50

Qy 68 HAYQILRKGGLKDENITIVMDIASSENPRPGVINKPDGEDYKGVPKDYTEKAVNV 127
Db 51 HAYQILFRHNGTPAEQIVVIMDIAVEDNTPTGIVINRPGTDYQGVFKDYTEDVT P 110

Qy 128 QNFYVNLGNGESVG-TG-GNGKVKGSGNDNIFTYADHGAGLJAMPTEGDEMAKDFNEV 186
Db 111 QNFYVLRGDAEAKVGIGSRKVLSKGQDHVFYFDHGSSGTLVFP-NEDLHVRLIKT 169

Qy 187 LEKMHKRKKYKNMVIYEACESGSMEFGLKLKNLNITYAATTAANSKESSWGVYCPEYPPP 246
Db 170 THYTFKRNMYRMKVFTIACSGSMRNH-LPDNINVYATTAANPRESSYACYYDE----- 223

Qy 247 PSEIGICLGDTFSISWEDSLHDMSKETLEQQHYVKRRVGSVPETSHVCRFGTEKML 306
Db 224 -KRSYLGDMVSNNMEDSDVEDLTKEUHQYHLVKSHT---NTSHVMQYGNKTIS 276

Qy 307 KDYLSSTIGRNPENDNFTFSFSS----PISNSGLVNPDIPLLYQRTIQAPGSL 361
Db 277 TMKVMQFQG ----MKRKASSPVPLPVTHLDLTPSPDVPLTIMKRLMT -NDL 325

Qy 362 ESK---EAQKLKDKEKNHRQOIDQSITDILRLSYKQTNVNLNSTRTGQPLYDDWDC 417
Db 326 EESRQLTEEYDYLARH-----LIRGEVEQ-----LLSERA--PLTGH-SC 364

Qy 418 FKTLYNSFEKHNHC 429
Db 365 YPEVLLYFRTHC 376

RESULT 4
US-09-193-524-2
Sequence 2, Application US/09193524
Patent No. 6007997
GENERAL INFORMATION:
APPLICANT: Braxton, Scott M.
APPLICANT: Delegane, Angelo M.
TITLE OF INVENTION: A NOVEL HUMAN CYSTEINE PROTEASE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCITE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94104

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/193, 524
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/928, 613
; FILING DATE: 08/928, 613
; APPLICATION NUMBER: 08/567, 506
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Luther, Barbara J.
; REGISTRATION NUMBER: 33954
; REFERENCE/DOCKET NUMBER: PF-0048 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-852-0195
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 431 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-193-524-2

Query Match 30.3%; Score 744; DB 3; Length 431;
Best Local Similarity 38.4%; Pred. No. 1.4e-64;
Matches 166; Conservative 70; Mismatches 128; Indels 68; Gaps 14;

Qy 8 FOLYLHALIFSAESRKTOQLLNDVSESSDKSAGTRWAVLAGSNEYNNYRHODIC 67
Db 3 WRVVVFSLVAIGA ----- VPIDPEDGKHHWVIVAGSGNWNYRQADAC 50

Qy 68 HAYQILRKGGLKDENITIVMDIASSENPRPGVINKPDGEDYKGVPKDYTEKAVNV 127
Db 51 HAYQILFRHNGTPAEQIVVIMDIAVEDNTPTGIVINRPGTDYQGVFKDYTEDVT P 110

Qy 128 QNFYVNLGNGESVG-TG-GNGKVKGSGNDNIFTYADHGAGLJAMPTEGDEMAKDFNEV 186
Db 111 QNFYVLRGDAEAKVGIGSRKVLSKGQDHVFYFDHGSSGTLVFP-NEDLHVRLIKT 169

Qy 187 LEKMHKRKKYKNMVIYEACESGSMEFGLKLKNLNITYAATTAANSKESSWGVYCPEYPPP 246
Db 170 THYTFKRNMYRMKVFTIACSGSMRNH-LPDNINVYATTAANPRESSYACYYDE----- 223

Qy 247 PSEIGICLGDTFSISWEDSLHDMSKETLEQQHYVKRRVGSVPETSHVCRFGTEKML 306
Db 224 -KRSYLGDMVSNNMEDSDVEDLTKEUHQYHLVKSHT---NTSHVMQYGNKTIS 276

Qy 307 KDYLSSTIGRNPENDNFTFSFSS----PISNSGLVNPDIPLLYQRTIQAPGSL 361
Db 277 TMKVMQFQG ----MKRKASSPVPLPVTHLDLTPSPDVPLTIMKRLMT -NDL 325

Qy 362 ESK---EAQKLKDKEKNHRQOIDQSITDILRLSYKQTNVNLNSTRTGQPLYDDWDC 417
Db 326 EESRQLTEEYDYLARH-----LIRGEVEQ-----LLSERA--PLTGH-SC 364

RESULT 5
US-09-134-001C-1248
Sequence 4248, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134, 001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064, 964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055, 779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 4248

RESULT 6
 US-09-453-702B-257
 ; Sequence 257, Application US/09453702B
 ; Patent No. 6365723
 ; GENERAL INFORMATION:
 ; APPLICANT: Blattner, Frederick R.
 ; Burland, Valerie
 ; Perna, Nicole T.
 ; Plunkett, Guy
 ; Welch, Rod
 ; TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
 ; NUMBER OF SEQUENCES: 265
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Quarles & Brady
 ; STREET: 1 South Pinckney Street
 ; CITY: Madison
 ; STATE: WI
 ; COUNTRY: US
 ; ZIP: 53701-2113
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Word Perfect 8.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/453-702B
 ; FILING DATE: 03-Dec-1999

CLASSIFICATION: <Unknown>
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 60/110, 955
 FILING DATE: 03-Dec-1998
 ATTORNEY/AGENT INFORMATION:
 NAME: Seay, Nicholas J.
 REGISTRATION NUMBER: 27386
 REFERENCE/DOCKET NUMBER: 960296.95017
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (608) 251-5000
 TELEFAX: (608) 251-9166
 INFORMATION FOR SEQ ID NO: 257:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3169
 TYPE: amino acid
 STRANDEDNESS: <Unknown>
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 SEQUENCE DESCRIPTION: SEQ ID NO: 257:

US-09-453-702B-257

Query Match Score 4.7% Best Local Similarity 21.2% Pred. No. 0.24; Length 3169;
 Matches 115; Conservative 83; Mismatches 188; Indels 157; Gaps 30;

Qy 18 LIFSAISRSRKTOOLLNDN--DEVESSDKSAK--GTRNAVLVAGSNE-VYNYRHOADITC 68
 Db 2496 1IYGNHGAKYTFLNNSGSKTIDITHLKDNSDYYDFRNLTIEHYTNELFISDNQGGF-- 2553

Query Match Score 115.5% Best Local Similarity 21.2% Pred. No. 0.24; Length 3169;
 Matches 115; Conservative 83; Mismatches 188; Indels 157; Gaps 30;

Qy 69 AYOILRKGGKLKDENTIVMDIAFSSENPRPGVINKPDGE---DVYKGYPKDTE 123
 Db 2554 VISULNNTAANINVRNMTSLOSS--GSILYIPLPSGDYHISDIYK MSRGKSF 2608

Qy 124 AVNQ----NFNVNLGNESGVTTGGNGKVKGSPN--DNIFYIYADHGAPGLIAMPTG 175
 Db 2609 KENVEKPKDIDIIINVAILETSYLQ----IKIPNNDDSDYLCLDR--PNLSS-- 2656

Qy 176 DEVMARDNEV-----LEMKHK-----RKYKMKMVYVEACESGSMFE 213
 Db 2657 ---YTLNENDLSGYISSLWDNIRGSFTPEHKTNTVNIAPNEKYTLISIGLDKLFNIDYFR 2713

Qy 214 GLLK-KNLNIYAVTAANSKSSWGVY-----CPCE-----SYPPPPSEIGTCUJD 256
 Db 2714 QALEVAKNKNSKVIS---KFTWETYGDIVVSPEDRISHLEDGFNYFSQP-ELDTPFSD 2767

Qy 257 TFSISM---LEDSLH---DMSKETLEQYH--VVKERRVSDVPENSVCRCFTEMML 306
 Db 2768 SFSYLYDNFOIQVDSVHKLHNRETKEOPIPHTRILKRYFIDSFAKTSITDR--EKNI 2824

Qy 307 -----KDYLISSYGRN-----PENINFTFESFSSPISNSGLVNPRLPIL 347
 Db 2825 YPVICDSDPDIHTSDYRHPFRIVLGNKTYPSPEELVKFLTSKEYLSNMNDVNNVIV-- 2882

Qy 348 YLQRKIQKAPGSLESKAOKLDEBKHNKRQIQTOSIMTILRUSVKQTNV----LN--- 399
 Db 2883 -----QKTTKKKNKLISVLSNSNIKIVLSCVMGTSKFLHLMNSG 2923

Qy 400 -LLTSTRTGQPLVDWDQCEKTVLVSF--KNHCGATVHYLGKVTGALANICNGDVVKQ 455
 Db 2924 DLLLTTSKTHGGVV--VIEKDFINNNWYNNLTLITPMGK1QE 2980

Qy 456 TYS 458
 Db 2981 TYS 2983

RESULT 7
 US-09-367-583-2
 ; Sequence 2, Application US/09367583
 ; Patent No. 6451566
 ; GENERAL INFORMATION:
 ; APPLICANT: Itoh, No. 6451566uya
 ; TITLE OF INVENTION: METHOD FOR PRODUCING DIHYDROXYACETONE-3-PHOSPHATE

```

; FILE REFERENCE: 06501/0388001
; CURRENT APPLICATION NUMBER: US/09/357,583
; EARLIER APPLICATION NUMBER: JP 1997-352610
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 2
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
; US-09-367-583-2

Query Match 4.7%; Score 114.5;
Best Local Similarity 19.3%; Pred. No. 0.
Matches 85; Conservative 72; Mismatched 11; LVCNCK

Qy 19 IFSAESRKTOOLLNDVESSDKSAKGTRAVLYAGS
Db 77 IFASPSRKRQIYTGIKQESE---AGT----LVICK

Qy 76 GGLKDENIIVMDIAF---SSENPRDG---VI
Db 123 AGKKAEE-LIAVADDSYGRKKSGKGVRGREGSISVY
Qy 127 ----VONFYNYVLLGNESGTGGNGKVYKG---PND
Db 177 IAKAAIDNLVSIASLAHVHVPCHEPIAKEDEMKHD
Qy 177 EYMAKDFNEVLEKMHKIRKVKYKNAVIVYFACEGSMF
Db 237 DLIAQMOKMLQDSDKDRAYVKI-----
Qy 237 VYCPESWPPPSEIGTCUGDTFTSISWLEDSDLHMS
Db 262 -----DDEVVJLMNNLGUSMLESAIASIKV
Qy 297 VCRFG-----TMLKRDYLSSYIGRNPDENDNT
Db 311 LNGLGFCTTLLRTDRV-----KVEGEYES
Qy 347 LYQRLKQKAPGSLESKEAOQKL-----LDEKNHR
Db 359 -----SKNKGNVSSEBQKVSKPVTVDKEKVR
Qy 402 TSTRTRGQLPYDDWDCKFL 421
Db 408 -----DTMAGBDGDCGTTL 420

RESULT 8
US-09-357-251-37
; Sequence 37, Application US/09357251
; Patient No. 6221441
; GENERAL INFORMATION
; APPLICANT: Falco, S. Carl
; APPLICANT: Fanodu, Layo O.
; APPLICANT: Orczco, Buddy
; APPLICANT: Schwaber, James S.
; TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthesis
; FILE REFERENCE: BB-1193
; CURRENT APPLICATION NUMBER: US/09/357,251
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Microsoft Office 97
; SEQ ID NO: 37
; LENGTH: 1440
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-357-251-37

```

Query Match	4.68;	Score	112.5;	DB	4;	Length	1440;
Best Local Similarity	22.8%	Pred. No.	0.13;				20;
Matches	87;	Conservative	56;	Mismatches	140;	Indels	99;
						Gaps	
Qy	51 VAGSNEYY ----- NYRHO ----- ADICHAYQILRKGGGLKDENTIVFMDIAFSSEN----- 98						
Db	562 VLGRDEDFFQVNNKNSKHELMGDPC --LKDLDKGDIOLQRGFFTCQDQEPEVSPYS 619						
Qy	99 ---RPGVILNKPDGEDVYKGVPKDYTEKAVNVQNFYNVLLNGESGVTTGGNGKVVKSGPND 155						
Db	620 CKEACPVLVYIPDGHT--KEMPTSGSKETKTKV-----ATKNET-----SAP-- 659						
Qy	156 NIFIYADGAPGLIAMPTEGDEMAKDENEV----LEKMHKRKKVNMKVIVTEACSGS 210						
Db	660 ----FKEQRTPSLNNNCTSESDLVLYNQVAVQDVVRLEKKAKKAKEDV----- 705						
Qy	211 MFEGLKLKNLNIYATAANSKESSWGYCPESYPPPSSEIGTCLGDTFSISWLEDSDLHD 270						
Db	706 -DAVKQLLSL---KAIFYKERTQGEYKPGN--PPAEQGNISSSSASLESKSLYD 756						
Qy	2711 MSKETLEQQYHVVYKRRVGSVPET----SHVCRFG----TEKMKDYLSSYGRNPEND 321						
Db	757 ----EVAAGEVVKRLKAEPSPKAKINAEVECULLSKAQKYKERTKEYIP--GQPPLSQ 809						
Qy	322 NFTFTESFSSPISMS---GLVNPRDPIYLQORKIQKAPMGSLESKEAQKKLDEKAHRK 378						
Db	810 S-----SDSSSPTRNSEPAGELET-----EAKVLFDFKVAS--QGEVVRKLKTEKAPD 854						
Qy	.379 QIDOSITDLRLSYKOTNWNLN 400						
Db	855 QVDIAVOELLOLKAQYKSILGV 876						
RESULT 9							
US-09-443-184-48							
; Sequence 48 Application US/09443184A							
; Patent No. 6372431							
; GENERAL INFORMATION:							
; APPLICANT: Cunningham, Mary Jane							
; APPLICANT: Zweizer, Gary							
; APPLICANT: Kaser, Matthew R.							
; APPLICANT: Panzer, Scott R.							
; APPLICANT: Sellhammer, Jeffrey J.							
; APPLICANT: Yue, Henry							
; APPLICANT: Baughn, Mariah							
; APPLICANT: Azimai, Yalda							
; APPLICANT: Lal, Preeti							
; TITLE OF INVENTION: MAMMALIAN TOXICOLOGICAL RESPONSE MARKERS							
; FILE REFERENCE: PC-0007 US							
; CURRENT APPLICATION NUMBER: US/09/443,184A							
; CURRENT FILING DATE: 1999-11-19							
; NUMBER OF SEQ ID NOS: 138							
; SOFTWARE: PERL program							
; SEQ ID NO: 48							
; LENGTH: 1512							
; TYPE: PRT							
; ORGANISM: Homo sapiens							
; FEATURE: misc_feature							
; OTHER INFORMATION: Incyte ID No. 6372431 2302721CD1							
US-09-443-184-48							
Query Match	4.68;	Score	112.5;	DB	4;	Length	1512;
Best Local Similarity	22.8%	Pred. No.	0.14;				20;
Matches	87;	Conservative	56;	Mismatches	140;	Indels	99;
						Gaps	
Qy	51 VAGSBEYY ----- NYRHO ----- ADICHAYQILRKGGGLKDENTIVFMDIAFSSEN----- 98						
Db	634 VLGRDEDFFQVNNKNSKHELMGDPC --LKDLDKGDIOLQRGFFTCQDQEPEVSPYS 691						
Qy	99 ---RPGVILNKPDGEDVYKGVPKDYTEKAVNVQNFYNVLLNGESGVTTGGNGKVVKSGPND 155						
Db	692 CKEACPVLVYIPDGHT--KEMPTSGSKETKTKV-----ATKNET-----SAP-- 731						

TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-471-033-50

Score 103; DB 1; Length 1338;
 Best Local Similarity 20.8%; Pred. No. 0.99; Mismatches 144; Indels 114; Gaps 19;
 Matches 83; Conservative 59;

Oy 81 ENITIVFMY --DDIAFSSNPENPRGVVINKPQGDVDYKGPDKYTKEAVINV ---QNFYNV 133
 Oy 343 ENITIVYRWCMPGPERFGYOISDPLRSL-----ADKEEQFLNTIKEDKCYMST 387

Db 134 LLGNESGVGTGGNGKVKSGPDNDNIFIYVADHGAPGLTAMP7GDEVMAKDFNEVLEKMHKR 193
 Qy 388 SLSERLAAFGSRI-----ILRLQVPKGSTGAYLSAIGFASEK --*IL--LDKD 434

Qy 194 KKYNRMVIVVEAESGSMSMFEGLK --KNUNIYAVTAANNSKESWMGYCPESYPPPSEIG 251
 Db 435 SKYH ----- IDKVTEVILKGVKRYVVDATLLNNSGSPSTPPTPSPSPPTPSD1G 484

Qy 252 TCGDGFSTSWLSDJHDMSKETLLEQQYHVVKRRVGSDDVPTSHYCRFG-----301
 Db 485 STM KTNOIStTQKNQKEMDRKGLLGYFKGK -----DFSNLTMFAPTRDSTLJYD 535

Qy 302 -----TEMMLKDYLS-YIG --RNPNDEIFT-----ESFSPFSNSG-----337
 Db 536 QQTANKLJQKQOQYQSRTWIGLJQSKEKGDEFNLSEDQDIEINGK1SNKGKEKV 595

Qy 338 --LYNPRDPLYLQORKIYQKAPGSLESKEAQKKLIDEEHKRQIDOSITDLR --LSV 392
 Db 596 VHLERKGKLVPIKIEQSDTFKFNDSKTPKELKFQDSQPOQVOO --DELRNPEFNK 652

Qy 393 KQT -----INVNLNT -----SFRTGQPLWDW 416
 Db 653 KESQEFPLAKPSKINFTQOMKREIDDTDGD5IPDLWE 692

RESULT 11
 US-08-471-044-50
 Sequence 50; Application US/08471033
 ; Patent No. 5770696

GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Koziel, Michael G
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Desai, Nalini M
 APPLICANT: Kosticichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 5770696el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/471,033
 FILING DATE: 23-MAR-1994
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/037,057
 FILING DATE: 25-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Pace, Gary M.
 REGISTRATION NUMBER: P-40,403
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8582
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1338 amino acids

RESULT 11
 US-08-471-044-50
 Sequence 50; Application US/08471044
 ; Patent No. 5840866
 GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Koziel, Michael G
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Desai, Nalini M
 APPLICANT: Kosticichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 5840868el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESS: CIBA-GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/471,044
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 800
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/463,483
 FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/314, 594
 FILING DATE: 09-SEP-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/218, 018
 FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/037, 057
 FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:
 NAME: Pace, Gary M.
 REGISTRATION NUMBER: 40,403
 REFERENCE/DOCKET NUMBER: CGC 1695/CIP3/DIV6 - SQLv3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8512
 TELEFAX: 919-541-8689

INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1338 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-471-04-50

Query Match 4.2%; Score 103; DB 2; Length 1338;
 Best Local Similarity 20.8%; Pred. No. 0.99; Mismatches 144; Indels 114; Gaps 19;
 Matches 83; Conservative 59; MisMatches 144; Indels 114; Gaps 19;

Qy 81 ENIVFMY---DDIAFSSENPRPGVYINKPDGEDVYKGVPKDTKTEAVNV---QNFYVN 133
 Db 343 ENIVTYRWCGMPFEFYQISDPLPSL-----KDFEEQFLNTIKEDKGYNST 387

Qy 134 LLGNESGVITGGNGKVVKSGPNDNIFTIYADHGAPGLIAMPTGDEVMADNEVLEKMKHR 193
 Db 388 SLSSERLAAGFSRKI-----ILRLQVPKGSAYLSSAIGGFASEK---EIL--LDKD 434

Qy 194 KKYNNKRMVIVYEACESGSMFFGILK -KNNLYIAYTAANSKESSMGVYCPESYPPPPSEIG 251
 Db 435 SKYH-----IDKVTTEVIKGYKRVYWDATLLNSRGPSPSTPPSPSTPPSPSDIG 484

Qy 252 TCLGDTFSISWLEDSDLHDSKETLEQQYHVVKRRVGSVDVPETSHVCREFG----- 301
 Db 485 STM-KTNQISTTQNQQKEMDRKGILGYFKGK-----DEFNLTMAPTRDSTLIYD 535

Qy 302 -----TEMLKDYL-SVIG--RNPDENDNFIT-----ESFSSPISNG---- 337

Db 536 QQTANKLKDQKOEYQSIRWKLGSKETSEAQQKILDEKNHRKOIDQSTIDLR--LSV 395

Qy 338 --LYNPRDIPIPYLORKIQKAPMGSLESKAQKFNLSEDEQATEILEINGKTIISNGKEKQY 595

Db 596 VHLERGKLKPVTKIEYQSDTKFNDSKTFKLKFIDSQNQPQVQQ--DEELANPEFNK 652

Qy 393 KQT-----NVNLNLT-----STRTGQPLVDDND 416
 Db 653 KESQEFLAKPSKINLFTOQMRETDDETDGDGSDIPDLWE 692

RESULT 12
 US-08-463-483A-50
 Sequence 50, Application US/08463483A
 Patent No. 5849870

GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Koziel, Michael G
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Desai, Nalini M
 APPLICANT: Kostichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:
 ADDRESSEE: CIBA GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NJ
 COUNTRY: USA
 ZIP: 10532

COMPUTER READABLE FORM:
 COMPUTER: IBM PC Compatible
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463, 483A
 FILING DATE:
 CLASSIFICATION: 530
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/314, 594
 FILING DATE: 09-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/218, 018
 FILING DATE: 23-MAR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/037, 057
 FILING DATE: 25-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Spruill, W. Murray
 REGISTRATION NUMBER: 32,943
 REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8815
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1338 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-463-483A-50

Query Match 4.2%; Score 103; DB 2; Length 1338;
 Best Local Similarity 20.8%; Pred. No. 0.99; Mismatches 144; Indels 114; Gaps 19;
 Matches 83; Conservative 59; MisMatches 144; Indels 114; Gaps 19;

Qy 81 ENIVFMY--DDIAFSSENPRPGVYINKPDGEDVYKGVPKDTKTEAVNV---QNFYVN 133
 Db 343 ENIVTYRWCGMPFEFYQISDPLPSL-----KDFEEQFLNTIKEDKGYNST 387

Qy 134 LLGNESGVITGGNGKVVKSGPNDNIFTIYADHGAPGLIAMPTGDEVMADNEVLEKMKHR 193
 Db 388 SLSSERLAAGFSRKI-----ILRLQVPKGSAYLSSAIGGFASEK---EIL--LDKD 434

Qy 194 KKYNNKRMVIVYEACESGSMFFGILK -KNNLYIAYTAANSKESSMGVYCPESYPPPPSEIG 251
 Db 435 SKYH-----IDKVTTEVIKGYKRVYWDATLLNSRGPSPSTPPSPSTPPSPSDIG 484

Qy 252 TCLGDTFSISWLEDSDLHDSKETLEQQYHVVKRRVGSVDVPETSHVCREFG----- 301
 Db 485 STM-KTNQISTTQNQQKEMDRKGILGYFKGK-----DEFNLTMAPTRDSTLIYD 535

Qy 302 -----TEMLKDYL-SVIG--RNPDENDNFIT-----ESFSSPISNG---- 337

Db 536 QQTANKLKDQKOEYQSIRWKLGSKETSEAQQKILDEKNHRKOIDQSTIDLR--LSV 395

Qy 338 --LYNPRDIPIPYLORKIQKAPMGSLESKAQKFNLSEDEQATEILEINGKTIISNGKEKQY 595

Db 596 VHLERGKLKPVTKIEYQSDTKFNDSKTFKLKFIDSQNQPQVQQ--DEELANPEFNK 652

Qy 393 KQT-----NVNLNLT-----STRTGQPLVDDND 416
 Db 653 KESQEFLAKPSKINLFTOQMRETDDETDGDGSDIPDLWE 692

RESULT 12
 US-08-463-483A-50
 Sequence 50, Application US/08463483A
 Patent No. 5849870

GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Koziel, Michael G
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Desai, Nalini M
 APPLICANT: Kostichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50

Query Match 4.2%; Score 103; DB 2; Length 1338;
 Best Local Similarity 20.8%; Pred. No. 0.99; Mismatches 144; Indels 114; Gaps 19;
 Matches 83; Conservative 59; MisMatches 144; Indels 114; Gaps 19;

Qy 81 ENIVFMY--DDIAFSSENPRPGVYINKPDGEDVYKGVPKDTKTEAVNV---QNFYVN 133
 Db 343 ENIVTYRWCGMPFEFYQISDPLPSL-----KDFEEQFLNTIKEDKGYNST 387

Qy 134 LLGNESGVITGGNGKVVKSGPNDNIFTIYADHGAPGLIAMPTGDEVMADNEVLEKMKHR 193
 Db 388 SLSSERLAAGFSRKI-----ILRLQVPKGSAYLSSAIGGFASEK---EIL--LDKD 434

Qy 194 KKYNNKRMVIVYEACESGSMFFGILK -KNNLYIAYTAANSKESSMGVYCPESYPPPPSEIG 251
 Db 435 SKYH-----IDKVTTEVIKGYKRVYWDATLLNSRGPSPSTPPSPSTPPSPSDIG 484

Qy 252 TCLGDTFSISWLEDSDLHDSKETLEQQYHVVKRRVGSVDVPETSHVCREFG----- 301
 Db 485 STM-KTNQISTTQNQQKEMDRKGILGYFKGK-----DEFNLTMAPTRDSTLIYD 535

Qy 302 -----TEMLKDYL-SVIG--RNPDENDNFIT-----ESFSSPISNG---- 337

Db 536 QQTANKLKDQKOEYQSIRWKLGSKETSEAQQKILDEKNHRKOIDQSTIDLR--LSV 395

Qy 338 --LYNPRDIPIPYLORKIQKAPMGSLESKAQKFNLSEDEQATEILEINGKTIISNGKEKQY 595

Db 596 VHLERGKLKPVTKIEYQSDTKFNDSKTFKLKFIDSQNQPQVQQ--DEELANPEFNK 652

Qy 393 KQT-----NVNLNLT-----STRTGQPLVDDND 416
 Db 653 KESQEFLAKPSKINLFTOQMRETDDETDGDGSDIPDLWE 692

RESULT 12
 US-08-463-483A-50
 Sequence 50, Application US/08463483A
 Patent No. 5849870

GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Koziel, Michael G
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Desai, Nalini M
 APPLICANT: Kostichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 5849870el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50

RESULT 13
US-08-471-046A-50 Application US/08471046A

Patent No. 5865346

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: Method For Isolating Vegetative Insecticidal Protein Genes

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: NO. 58663126artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

ZIP: 27709

COMPUTER READABLE FORM:

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/471,046A
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/463,483
FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Melgs, J. J. Timothy
REGISTRATION NUMBER: 38 241

REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV8 - SQLv4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

SEQUENCE CHARACTERISTICS:

LENGTH: 1338 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match Score 103; DB 2; Length 1338;
Best Local Similarity 20.8%; Prod. No. 0.99;
Matches 83; Conservative 59; Mismatches 144; Indels 114; Gaps 19;

Qy 81 ENITIVFM--DIAFSSENPPGVINPKDGEVKGPKDYTPEAVN--QNFNYV 133
Db 343 ENTVYRCCGMPFGQYQISDPLSL-----KDFEFOFLNTIKERGYMST 387

Qy 134 LIGNESSGTGGKVVSGPNDNIFIYYADHGAPGLJAMPGDEYMAKDFNEVLEKMKR 193
Db 388 SLSERLAAFGSJKI----ILRLOVKGSTGAYLSAIGFASEK--EIL-LDKD 434

Qy 194 KYNKMYIVYACEGSMFEGILK - KNLNLYATAANSKESSNGVYCPESYPPPSEIG 251.

RESULT 14
US-08-470-566B-50 Application US/08470566B

Patent No. 5872212

GENERAL INFORMATION:

APPLICANT: Warren, Gregory W
APPLICANT: Koziel, Michael G
APPLICANT: Mullins, Martha A
APPLICANT: Nye, Gordon J
APPLICANT: Carr, Brian
APPLICANT: Desai, Nalini M
APPLICANT: Kostichka, N. Kristy
APPLICANT: Duck, Nicholas B
APPLICANT: Estruch, Juan J
TITLE OF INVENTION: No. 5872212ei. Pesticidal Proteins and Strains
NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: No. 5872212artis Corporation

STREET: 3054 Cornwallis Road

CITY: Research Triangle Park

STATE: NC

COUNTRY: USA

ZIP: 27709

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30B

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/470,566B
FILING DATE: 06-JUN-1995
CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/463,483
FILING DATE: 05-JUN-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/314,594
FILING DATE: 09-SEP-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/218,018
FILING DATE: 23-MAR-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/037,057
FILING DATE: 25-MAR-1993

ATTORNEY/AGENT INFORMATION:

NAME: Melgs, J. J. Timothy
REGISTRATION NUMBER: 38 241

REFERENCE/DOCKET NUMBER: CGC1695/CIP3/DIV4

TELECOMMUNICATION INFORMATION:

TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689

SEQUENCE CHARACTERISTICS:

LENGTH: 1338 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match Score 103; DB 2; Length 1338;
Best Local Similarity 20.8%; Prod. No. 0.99;
Matches 83; Conservative 59; Mismatches 144; Indels 114; Gaps 19;

Qy 81 ENITIVFM--DIAFSSENPPGVINPKDGEVKGPKDYTPEAVN--QNFNYV 133
Db 343 ENTVYRCCGMPFGQYQISDPLSL-----KDFEFOFLNTIKERGYMST 387

Qy 134 LIGNESSGTGGKVVSGPNDNIFIYYADHGAPGLJAMPGDEYMAKDFNEVLEKMKR 193
Db 388 SLSERLAAFGSJKI----ILRLOVKGSTGAYLSAIGFASEK--EIL-LDKD 434

Qy 194 KYNKMYIVYACEGSMFEGILK - KNLNLYATAANSKESSNGVYCPESYPPPSEIG 251.

SEQUENCE CHARACTERISTICS:

LENGTH: 1338 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-470-566B-50

Query Match Score 103; DB 2; Length 1338;
 Best Local Similarity 20.8%; Pred. No. 0.99; Gaps 19;
 Matches 83; Conservative 59; Mismatches 144; Indels 114; Gaps 19;

Qy 81 ENIVFMY---DDIAFSSENPRPGVINKPDKDGYKGVPKDYTEKEAVNV---QNFYNY 133
 Db 343 ENITYVRWCGMPEFYQISDPLSL-----KDFEEQFLNTIKEDKGYMT 387

Qy 134 LLGNESGVTTGNGKVVKSGPNDNIFTLYADHGAPSLIAMPDTGDEVMAKDENELEKMKHR 193
 Db 388 SLSSEERLAAGRSRKI-----ILRLQPKGSIGAYLSAIGGFASEK---EIL-LDKD 434

Qy 194 KYNKRMVIVYEACESGSMFFEGIL-KNLYNIAYTAANSKESWGVCPESYPPPSEIG 251
 Db 435 SKYH-----IDKVTEVILIGKVKRYVVDATLTLNSRGSPSTPPSSTPTPSDIDG 484

Qy 252 TCLGDTFSISWLEDSDLHMSKETLEQQYHVKKRRVSDVPETSHVCRCFG-----301
 Db 485 STM-KTNOIISTTKNQQKEMDRKGLLGYFKGK-----DEFNLTMFAPTRDSTLIYD 535

Qy 302 -----TEKMLKDYL-SVIG--RNPNENDNFTT-----ESFSSPINSG----337
 Db 536 QQTANKLKDQKQEYQSTIRNMIGLQSKETGDFTEFLSEQAITEINGKISNGKKEQV 595

Qy 338 --LYNPRDITPLLYLORKIQKAPMGSLESKPAQKULDEKNHRKOIDQSITDILR--LSV 392
 Db 596 VHLERGKLVPKIEQSDTKFNIDSKTFKELKFIDSQNQQVQQ--DELNNEPEFNK 652

Qy 393 KQT-----NVNLNT-----STRTGQPLVDDND 416
 Db 653 KESQEFFLAKPSKINLFTQMKREIDEDTDIDGSTPDLWE 692

RESULT 15
 US-08-469-334-50
 Sequence 50, Application US/08469334
 Patent No. 5903B3

GENERAL INFORMATION:
 APPLICANT: Warren, Gregory W
 APPLICANT: Mullins, Martha A
 APPLICANT: Nye, Gordon J
 APPLICANT: Carr, Brian
 APPLICANT: Desai, Nalini M
 APPLICANT: Kostichka, N. Kristy
 APPLICANT: Duck, Nicholas B
 APPLICANT: Estruch, Juan J
 TITLE OF INVENTION: No. 590383el Pesticidal Proteins and Strains
 NUMBER OF SEQUENCES: 50
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CIBA-GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: NY
 COUNTRY: USA
 ZIP: 10532

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30B
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/469,334
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 800
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 08/463,483

FILING DATE: US 08/314,594
 APPLICATION NUMBER: 09-REP-1994
 FILING DATE: 09-REP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/218,018
 FILING DATE: 23-MAR-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/037,057
 FILING DATE: 25-MAR-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Spaulill, W. Murray
 REGISTRATION NUMBER: 32,943
 REFERENCE/DOCKET NUMBER: CGC 1695/CIP3
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 919-541-8615
 TELEFAX: 919-541-8689
 INFORMATION FOR SEQ ID NO: 50:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1338 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-469-334-50

Query Match Score 103; DB 2; Length 1338;
 Best Local Similarity 20.8%; Pred. No. 0.99; Gaps 19;
 Matches 83; Conservative 59; Mismatches 144; Indels 114; Gaps 19;

Qy 81 ENIVFMY--DDIAFSSENPRPGVINKPDKDGYKGVPKDYTEKEAVNV--QNFYNY 133
 Db 343 ENITYVRWCGMPEFYQISDPLSL-----KDFEEQFLNTIKEDKGYMT 387
 Db 388 SLSSEERLAAGRSRKI-----ILRLQPKGSIGAYLSAIGGFASEK---EIL-LDKD 434
 Db 388 SLSSEERLAAGRSRKI-----ILRLQPKGSIGAYLSAIGGFASEK---EIL-LDKD 434
 Qy 194 KYNKRMVIVYEACESGSMFFEGIL-KNLYNIAYTAANSKESWGVCPESYPPPSEIG 251
 Db 435 SKYH-----IDKVTEVILIGKVKRYVVDATLTLNSRGSPSTPPSSTPTPSDIDG 484
 Db 435 SKYH-----IDKVTEVILIGKVKRYVVDATLTLNSRGSPSTPPSSTPTPSDIDG 484
 Qy 252 TCLGDTFSISWLEDSDLHMSKETLEQQYHVKKRRVSDVPETSHVCRCFG-----301
 Db 485 STM-KTNOIISTTKNQQKEMDRKGLLGYFKGK-----DEFNLTMFAPTRDSTLIYD 535
 Db 536 QQTANKLKDQKQEYQSTIRNMIGLQSKETGDFTEFLSEQAITEINGKISNGKKEQV 595
 Qy 302 -----TERMLKDYL-SVIG--RNPNENDNFTT-----ESFSSPINSG----337
 Db 536 QQTANKLKDQKQEYQSTIRNMIGLQSKETGDFTEFLSEQAITEINGKISNGKKEQV 595
 Qy 338 --LYNPRDITPLLYLORKIQKAPMGSLESKPAQKULDEKNHRKOIDQSITDILR--LSV 392
 Db 596 VHLERGKLVPKIEQSDTKFNIDSKTFKELKFIDSQNQQVQQ--DELNNEPEFNK 652
 Qy 393 KQT-----NVNLNT-----STRTGQPLVDDND 416
 Db 653 KESQEFFLAKPSKINLFTQMKREIDEDTDIDGSTPDLWE 692

Search completed: May 27, 2003, 15:29:06
 Job time : 31 secs

THIS PAGE BLANK (USPTO)

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model.

Run On: May 27, 2003, 15:27:45 ; Search time 57 Seconds (without alignments)

Scoring table: BLOSUM62 810.780 Million cell updates/sec

Title: US-09-934-066-2

Perfect score: 2454

Sequence: MSSPAGHQLVFLHALLIF.....CNMGVDVKOTVSAIEQACSM 4 666

Post-processing: Minimum Match 0% Maximum Match 100%

Scoring table: Gapext 10.0 , Gapext 0.5

Searched: 375593 seqs, 99172665 residues

Total number of hits satisfying chosen parameters: 375593

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:
 1: /cn2_6/podata/2/pubpa/US08_NEW_PUB.pep:
 2: /cn2_6/podata/2/pubpa/PCT_NEW_PUB.pep:
 3: /cn2_6/podata/2/pubpa/US06_NEWPUB.pep:
 4: /cn2_6/podata/2/pubpa/US06_NEWPUB.pep:
 5: /cn2_6/podata/2/pubpa/US07_NEWPUB.pep:
 6: /cn2_6/podata/2/pubpa/US07_NEWPUB.pep:
 7: /cn2_6/podata/2/pubpa/PECTUS_PUBCOMB.pep:
 8: /cn2_6/podata/2/pubpa/US08_NEWPUB.pep:
 9: /cn2_6/podata/2/pubpa/US09_NEWPUB.pep:
 10: /cn2_6/podata/2/pubpa/US09_NEWPUB.pep:
 11: /cn2_6/podata/2/pubpa/US10_NEWPUB.pep:
 12: /cn2_6/podata/2/pubpa/US10_NEWPUB.pep:
 13: /cn2_6/podata/2/pubpa/US60_NEWPUB.pep:
 14: /cn2_6/podata/2/pubpa/US60_NEWPUB.pep:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

```

RESULT
US-09-934-066-2
; Sequence 2, Application US/09334066
; Patent No. US20020108149A1
; GENERAL INFORMATION:
; APPLICANT: Gruis, Darren B.
; INVENTION: Jung, Rudolf
; TITLE OF INVENTION: Methods of Increasing Polypeptide
; TITLE OF INVENTION: Accumulation in Plants
; FILE REFERENCE: 35718/237251
; CURRENT APPLICATION NUMBER: US/09/934,066
; CURRENT FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,804
; PRIOR FILING DATE: 2000-08-21
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; US-09-934-066-2

```

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2454	100.0	466	10	Sequence 2, Appli
2	744	30.3	431	9	Sequence 2, Appli
3	487	19.8	319	9	Sequence 2, Appli
4	115.5	4.7	3169	9	Sequence 257, Appli
5	114.5	4.7	580	9	Sequence 2, Appli
6	107	4.4	996	9	Sequence 3847, Appli
7	103	4.2	568	10	Sequence 514.0, Appli
8	102	4.2	1034	9	Sequence 43, Appli
9	102	4.2	1189	9	Sequence 35, Appli
10	100.5	4.1	2494	10	Sequence 3, Appli
11	100	4.1	828	10	Sequence 4, Appli
12	99	4.0	707	9	Sequence 313, Appli
13	99	4.0	707	10	Sequence 313, Appli
14	98.5	4.0	969	8	Sequence 3, Appli
15	98.5	4.0	26926	9	Sequence 2, Appli
16	97	4.0	503	10	Sequence 8, Appli
17	96.5	3.9	452	9	Sequence 255, Appli
18	96.5	3.9	452	9	Sequence 255, Appli
19	96.5	3.9	452	9	Sequence 255, Appli

Qy 241 ESTYPPPSEIGTCLGDFTSISNLEDSDLHDMSKETLEQOYHVKRKGSDVPEITSHVCRF 300
Db 241 ESTYPPPSEIGTCLGDFTSISNLEDSDLHDMSKETLEQOYHVKRKGSDVPEITSHVCRF 300
Qy 301 GTEKMLKDLYSSYIGRNPDENDFTESFSSPINSGLVNPRDIPPLYLQRKIQKAPMGS 360
Db 301 GTEKMLKDLYSSYIGRNPDENDFTESFSSPINSGLVNPRDIPPLYLQRKIQKAPMGS 360
Qy 361 LESKEAQKLDEKNRKQDOSITDLRLSVKQTNVNLLTSTRTGQPLVDMDCFKT 420
Db 361 LESKEAQKLDEKNRKQDOSITDLRLSVKQTNVNLLTSTRTGQPLVDMDCFKT 420
Qy 421 LVNSFKNHCGATVHYGILKYTGALANICNMGYDKVQTYSAEQACSM 466
Db 421 LVNSFKNHCGATVHYGILKYTGALANICNMGYDKVQTYSAEQACSM 466

RESULT 2
: Sequence 2, Application US/09967792
: Patent No. US200215535A1
: GENERAL INFORMATION:
: APPLICANT: Diep, Dinh
: Braxton, Scott M.
: Delegene, Angelo M.

TITLE OF INVENTION: A NOVEL HUMAN CYSTEINE PROTEASE
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/967,796
FILING DATE: 28-Sep-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/449,422
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Luther, Barbara J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0048 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 431 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-967-796-2

Qy 128 QNFIVNLGENESCVTG-GNGKVKSGPNDNPIYYADHGAGLAMPTEGVMDVAKDFNEV 186
Db 111 QNFIVLVLGDAEAVKGCSRKVLKGQDHVTFIDHGSTGILVFP-NEDLHVQDLSIKT 169
Qy 187 LEKMKRKKYKMKVTVYEACESGSMFEGILKKNNIYAVTAANSKESSEGVYCPESYPP 246
Db 170 THYIFKMKYKRVFYEAECGSMMH-LPONINVYATTAAAPRESSYACYYDB----- 223
Qy 247 PSEIGTCLGDFTSISNLEDSDLHDMSKETLEQOYHVKRKGSDVPEITSHVCRFTEKML 306
Db 224 --KRSTYLGDWTSVNWMEDSDVEDLTAKTLQHVLVKSHT----NTSHVMOYGNKTIS 276
Qy 307 KDISSYIGRNPDENDFTESFSS-----PTNSGLVNPRDIPPLYLQRKIQKAPMGS 361
Db 277 TMKVMQFOG-----MKRASSPPPLPVTHLDTPSPDVPLVTIMKRLMLMT-NDL 325
Qy 362 ESK----BAQKLLDEKNRKQDOSITDLRLSVKQTNVNLLTSTRTGQPLVDMDC 417
Db 326 EESRQLTEEQYLDARH-----LIREVEQ-----LLSERA--PLTGH-SC 364
Qy 418 FKTLVNSFSKNC 429
Db 365 YPEVLLYFRTHC 376

RESULT 3
US-10-043-487-298
: Sequence 298, Application US/10043487
: Publication No. US2003055220A1
: GENERAL INFORMATION:
: APPLICANT: Pierre, LEGRAIN
: TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides
: FILE REFERENCE: B4778A
: CURRENT APPLICATION NUMBER: US/10/043-487
: PRIOR APPLICATION NUMBER: US 60/261,130
: NUMBER OF SEQ ID NO: 561
: SEQ ID NO: 298
: LENGTH: 319
: TYPE: PRT
: ORGANISM: Shigella Flexneri
US-10-043-487-298

Query Match 19.8%; Score 487; DB 9; Length 319;
Best Local Similarity 33.5%; Pred. No. 1e-33; Indels 124; Gaps 12;
Matches 115; Conservative 64; Mismatches 124;

Qy 133 VLIGNESSCVTG-GNGKVKSGPNDNPIYYADHGAGLAMPTEGVMDVAKDFNEVLEKMH 191
Db 2 VLRGDAEAVKGCSRKVLKGQDHVTFIDHGSTGILVFP-NEDLHVQDLSIKT 60
Qy 192 KRMKVKMKVTVYEACESGSMFEGILKKNNIYAVTAANSKESSEGVYCPESYPP 251
Db 61 KHMVKRKYKRVFYEAECGSMMH-LPONINVYATTAAAPRESSYACYYDB-----KRS 112
Qy 252 TCLGDFTSISNLEDSDLHDMSKETLEQOYHVKRKGSDVPEITSHVCRFTEKMLQDLS 311
Db 113 TYLGDWTSVNWMEDSDVEDLTAKTLQHVLVKSHT----NTSHVMOYGNKTIS 167
Qy 312 SYIGRNPDENDFTESFSS-----PTNSGLVNPRDIPPLYLQRKIQKAPMGS-ESKE 365
Db 168 QFG-----MKRASSPPPLPVTHLDTPSPDVPLVTIMKRLMLMT-NDLESRSQ 216
Qy 366 AOKLLDEKNRKQDOSITDLRLSVKQTNVNLLTSTRTGQPLVDMDCFKTLVNSF 425
Db 217 LTERIORHLDAHLRSVRKIVSLLAASEAEVEQLSERA--PLTGH-SCYPEALLHF 272
Qy 426 KNHC---GATVHYGLKTYGALANICNMGYDKVQTYSAEAC 464

Query Match 30.3%; Score 744; DB 9; Length 431;
Best Local Similarity 38.4%; Pred. No. 1.3e-55; Indels 68; Gaps 14;
Matches 166; Conservative 70; Mismatches 128;

Qy 8 FOILVFLHALIFSABSRKTQOLLNDVSESSDKSAKGTRNAVLYAGSNEXYNYRHOADIC 67
Db 3 WVVVFISVALIGA-----VPIDDPEOGKHHVVIAGSNWNYTHOADC 50
Qy 68 HAYQILRKGGLKDENITIVMFDDIAFSSENMPRGVYLINKPDGEVYKGPKDYTEKAV/N 127
Db 51 HAYQFIRNGTFAEQIVIMYDDIASEDNTPGIVNPNGTIDYQGVPKDFTGEDVTP 110

RESULT 4
US-10-114-170-257

Sequence 257 Application US/10114170
Publication No. US2010023075A1

GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.

Plunkett, Guy
Welch, Rod

TITLE OF INVENTION: No. US2010023075A1 Sequences of E. coli O157

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:
ADDRESSEE: Quarles & Brady

STREET: 1 South Pinckney Street

CITY: Madison

STATE: WI

COUNTRY: US

ZIP: 53701-2113

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 8.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US10/114,170

FILING DATE: 01-APR-2002

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/453,702

FILING DATE: 03-DEC-1999

APPLICATION NUMBER: 60/110,955

FILING DATE: 04-DEC-1998

ATTORNEY/AGENT INFORMATION:

NAME: Seay, Nicholas J.

REGISTRATION NUMBER: 27,386

REFERENCE/DOCKET NUMBER: 960296-95017

TELECOMMUNICATION INFORMATION:

TELEPHONE: (608) 251-2000

TELEFAX: (608) 231-9166

INFORMATION FOR SEQ ID NO: 257:

SEQUENCE CHARACTERISTICS:

LENGTH: 3169 amino acids

TYPE: amino acid

STRANDEDNESS: <Unknown>

TOPOLOGY: Linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 257:

US-10-114-170-257

Query Match 4.7% Score 115.5; DB 9; Length 3169;
Best Local Similarity 21.2%; Pred. No. 2.0; Gaps
Matches 115; Conservative 83; Mismatches 188; Indels 157; Gaps

Qy 18 LIFSAEFSKTOQLNDN---DVESSDKSAK---GTRWAVLVAGSN--YVNYRHQADICH 68
Db 2496 IYGNHGAKYTFLNSSGKTDIHLKNSDIDTFDFRNNTIIFHYTNELFISFDNQGGF-- 2553

Qy 69 AYQILRKGLKDENTIVMFYDDIAFSSENPRGVLTINKPDGE---DVKYGPDKDVTKE 123
Db 2554 VISILNNATSEANINVFRKNNTSLOSS---GSLTVLPSGDIVYHISDLYK-MSRGRKSF 2608

Qy 124 AVNQV----NFYNVNLGNESGTGGNGKVKSGPN--DNIFIYADHGAPLIAIMTG 175
Db 2609 KLNVEKKPDDIIINVAILTETSYLQ----IKKIPNNDSDYILCLDN--PNLSS---- 2656

Qy 176 DEVMKADEVY-----LEKMRHK-----RKYKNMKTYVEAGEGSQSMFE 213
Db 2657 ---TINFNDSLGYISSLWQNIRGSETPEFHNTVNIAPEKKYISLLGLDKLSFNIDVFR 2713

Qy	214	GILK-KUNLNYIAYTAANSKESSWGVY-----CPE-----SYPPPSEIGTGLGD 256	
Db	2714	QALEVKKNNKSYKIS----KFTWETYGDIVVSPEDRISHLEDFNNSQOP ELDPIFSD 2767	
Qy	257	TFSLIW----LEDSLH----DMSKETLQQYH--VVKRRGSDVPTESHYCREFTEKML 306	
Db	2768	SFSYLYDNFOIVOSDVHITKLHLNREPKITHRIILKRYFDSFATISITDTR---EKNI 2824	
Qy	307	-KDYLSYYIGN-----PENDNFTTESSSPISNSGLYNPRDIPLL 347	
Db	2825	YPVCDSPDHFTEDYRPERTYLGNKNTLPSEELYEPISKEYLSNMDYINNVTP-- 2882	
Qy	348	YLQRKIQKAPMGSLSEAQKQLLDEKNHRKOIDQSSTDILRLSVKOTINV----LN--- 399	
Db	2883	-QKTTKKNNKLSTVLSNSNKNDIVSGNTGTGSKIFHNLNSNG 2923	
Qy	400	--LLTSTRTTGOLPDMDCFKTLVNSF--KHNCGATVHYGHLYKTGALANICNMGYDKO 455	
Db	2924	DLLLTTSKTHGGGVW--VIFKDFINNNWWKYNLTLLTVPIDNKLSDLNRINTPMGKIQE 2980	
Qy	456	TYS 458	
Db	2981	TYS 2983	
RESULT 5			
US-10-202-211-2			
; Sequence 2, Application US/10202211			
; Publication No. US20030030551A1			
; GENERAL INFORMATION:			
; APPLICANT: Itch, No. US20030030551auya			
; TITLE OF INVENTION: METHOD FOR PRODUCING DIHYDROXYACETONE-3-PHOSPHATE			
; FILE REFERENCE: 06501/038001			
; CURRENT APPLICATION NUMBER: US/10/202/211			
; CURRENT FILING DATE: 2002-07-22			
; PRIOR APPLICATION NUMBER: US/09/367,583			
; PRIOR FILING DATE: 1999-12-02			
; NUMBER OF SEQ ID NOs: 4			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 2			
; LENGTH: 580			
; TYPE: PRT			
; ORGANISM: Schizosaccharomyces pombe			
US-10-202-211-2			
Query Match 4.7% Score 114.5; DB 9; Length 580;			
Best Local Similarity 19.3%; Pred. No. 0..23; Mismatches 150; Indels 133; Gaps 2			
Matches 85; Conservative 72; Mismatches 150; Indels 133; Gaps 2			
Qy	19	IFSAESRKTLQIINNDVSSDKSAKGTWRAVLVAGSNEYNNYRHADITCH--AYOFLRK 75	
Db	77	IFASPSSKOIYTGIKQVESE----AGT----LVICKN----YTGDILHFGMALEKQRT 122	
Qy	76	GGLKDENIIVFMDIAF----SSENPRPG---VINKPDGEDVYGVPRDYTEKAVN- 126	
Db	123	AGKKAEL--LIAVADDVSYGRKKSKGVRGLSGTVLHKIAAAAANGVTP---LEAVTT 176	
Qy	127	----VQNPVNVLGNNEGGVTGGNGKVKRG--PNDNFIITYADHGPG--LIAMPTGD 176	
Db	177	IKAIAIDNLVIASLAVHVHPGHEP_LAKEDEMKHDMELGNGIHNPFGCKRISPITSID 236	
Qy	177	EVMKADFNEVLEKMKHKKYKVNMYIYEACESGSMFEGILKKNLNIYAVTAANSKESSWG 236	
Db	237	DLIAQMQLQMDSDKDRAYVKI-----EG-----	
Query Match 4.7% Score 114.5; DB 9; Length 580;			
Best Local Similarity 19.3%; Pred. No. 0..23; Mismatches 150; Indels 133; Gaps 2			
Matches 85; Conservative 72; Mismatches 150; Indels 133; Gaps 2			
Qy	237	VYCPESYPPPSEIGTCLGDTFISIWLLEDSDHMSKTELQOYHVKKRPGSDVPETSH 296	
Db	262	-----DDEVVLMLNNLGLSMLMEFAISHKYREALAKEYKINPVRIFAG-PTTS 310	
Qy	297	VCREFG-----TEKMLKDYLSYIYGRNPENDNFTTESESFSPINSG--LYNPNDIPL 346	
Db	311	LNGLGFITLRLTRDTRV-----KVEGEYSVLNLQDQPEAIGWPLOQPSDJK- 358	

RESULT 6
 Qy 347 LYLRQKIQAPMGSLESKEAQKKL-----LDEKKHQRQIDQSITIDILRLSVKOTINVNLN 401
 ; Sequence 3847, Application US/09738626
 ; Publication No. US20020197605A1
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAGAWA, SATOSHI
 ; APPLICANT: MIZOGUCHI, HIROSHI
 ; APPLICANT: ANDO, SEIKO
 ; APPLICANT: HAYASHI, MIKIRO
 ; APPLICANT: OCHIAI, KEIKO
 ; APPLICANT: YOKOI, HARUHIKO
 ; APPLICANT: TATEISHI, NAOKO
 ; APPLICANT: SENOH, AKIHIRO
 ; APPLICANT: IKEDA, MASATO
 ; APPLICANT: OZAKI, AKIO
 ; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
 ; CURRENT APPLICATION NUMBER: US/09/738,626
 ; FILE REFERENCE: 249-125
 ; CURRENT FILING DATE: 2000-12-18
 ; PRIORITY NUMBER: JP 99-377484
 ; PRIOR FILING DATE: 1999-12-16
 ; PRIOR APPLICATION NUMBER: JP 00/159162
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: JP 00/280988
 ; PRIOR FILING DATE: 2000-08-03
 ; NUMBER OF SEQ ID NOS: 7059
 ; SOFTWARE: Patentin ver. 3.0
 ; SEQ ID NO: 3847
 ; LENGTH: 996
 ; TYPE: PRT
 ; ORGANISM: Corynebacterium glutamicum
 ; US-09-738-626-3847

Query Match 4.4%; Score 107; DB: 9; Length 996;
 Best Local Similarity 19.9%; Pred. No. 2.3; Indels 193; Gaps 19;

Qy 23 ESRKTQLNDNEYESS----DKSARG----TRAVLVAGSNNEYNNRHOADICHAYQIL 73
 Db 565 ENFTSFLV-DPDTSSMEEDELDNAAAGREGRTEN----NGYFGDAAQDSMAEVA 617
 Qy 74 RKGGLK---DENI-----IVMFDIAFSSENPRP-----VIINKFGEDDY 113
 Db 618 RQGKLKAVALDANELEHDARSVNSNLKFDDAEGRANVRGRRGYTERIVGTAAEGEPEF 677
 Qy 114 K----GVPKDYTKAEVNVQ----NFYVNLGNESGVGGNGKVV-----KSGPNNDNIFIY 162
 Db 678 QRNALPBTTPDPLTLEVAERLFLATPOGGREBLGTTNPANGRMVVAEREGRGP--YVIEOVT 735
 Qy 163 DHGAPGLJAMPTEDEWMA-----KDFNEVLEKMHKREKYNKMYIYVEACIESG 205
 Db 736 DSERAG--AEOAEVVAEAERKADEFQRTGMRPKNWETKTAANQKEKRINQIVEENLK 793
 Qy 206 CEGSMSEGTEILKLNLYAATANSKESSNGVYCYESPPPSSEIGTCUGDTFISWIED 265
 Db 794 PATASLSGMEPATVTLLEALKLSSLPREVG-----DPSDNNEVITAQNRYGPYLKG 847
 Qy 266 SDLHDMSKETLEQYHVVKRPGSDVPTSHCRGTERMKLKVLSSTYGRNPENDNPF 325
 Db 848 SDERSLNE-EQKIFTVLTDEARAYIAEPKRRGRAAAQPPLKQGDNDYSGKP----MTV 901
 Qy 326 TESFSSPISNSGLVNPR---DIP-----LLYLRQKIQKAPMGSLESKEAQ---- 367

RESULT 7
 US-09-815-242-5140
 ; Sequence 5140, Application US/09815242
 ; Patent No. US20020061565A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Haselbeck, Robert
 ; APPLICANT: Ohlsen, Karl L.
 ; APPLICANT: Zyskind, Judith W.
 ; APPLICANT: Wall, Daniel
 ; APPLICANT: Trwisch, John D.
 ; APPLICANT: Carr, Grant J.
 ; APPLICANT: Yamamoto, Robert T.
 ; APPLICANT: Xu, H. Howard
 ; TITLE OF INVENTION: Identification of Essential Genes In
 ; PROkaryotes
 ; FILE REFERENCE: ELTRA_01A
 ; CURRENT APPLICATION NUMBER: US/09/815,242
 ; CURRENT FILING DATE: 2001-03-21
 ; PRIOR APPLICATION NUMBER: 60/191,078
 ; PRIOR FILING DATE: 2000-03-21
 ; PRIOR APPLICATION NUMBER: 60/206,848
 ; PRIOR FILING DATE: 2000-05-13
 ; PRIOR APPLICATION NUMBER: 60/207,727
 ; PRIOR FILING DATE: 2000-05-26
 ; PRIOR APPLICATION NUMBER: 60/242,578
 ; PRIOR FILING DATE: 2000-10-23
 ; PRIOR APPLICATION NUMBER: 60/253,625
 ; PRIOR FILING DATE: 2000-11-27
 ; PRIOR APPLICATION NUMBER: 60/257,931
 ; PRIOR FILING DATE: 2000-12-22
 ; PRIOR APPLICATION NUMBER: 60/269,308
 ; PRIOR FILING DATE: 2001-02-16
 ; NUMBER OF SEQ ID NOS: 14110
 ; SOFTWARE: FastSEQ For Windows Version 4.0
 ; SEQ ID NO: 5140
 ; LENGTH: 568
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 ; US-09-815-242-5140

Query Match 4.2%; Score 103; DB: 10; Length 568;
 Best Local Similarity 21.1%; Pred. No. 2.2; Indels 108; Gaps 16;
 Matches 65; Conservative 36; Mismatches 99; Score 103; DB: 10; Length 568;
 Qy 39 DRSGATPRWAVLVAGSMNEYNNRHOADICHAYQILRKGGURDENTIVEMYDIAFSEN 98
 Db 62 DODAK----LLVPFDNYD-----
 Qy 99 RFGVYINKPDGEDVYKGPKDYTKAEVNVN----QNFYVNLLGNESGVGGNCKVVKSG---- 152
 Db 94 RWNLIQK--DGK-----VLRDGEAISLLPMLYGFVRDQKYGVDGOCHEAOAPRFDD 147
 Qy 153 --PND-NIFITYADHGAPGLJAMPTEDEWMAKDFNEVLEKMHKREKYNKMYIYVEACIESG 209
 Db 148 IYPNSANEFITEIDGKRGIL----DAKCKQLTDEA-----YDFTLVNGSVAENG 193
 Qy 210 SMPEGILKRNLYAATANSKESSNGVYCYESPPPSSEIGTCUGDTFISWIEDSLDH 269
 Db 194 GLI-----SAERGEEKW-----LNLATGEQKAVAYESLGDLH 226
 Qy 270 D--MSKETLEQYHVVKRPGSDVPTSHCRGTERMKLKVLSSTYGRNPENDNPF 324
 Db 227 DGMSASVIGQSOLDAKGIVVGQDGDG-----KSY--DYLG-TPANGLYA 267
 Qy 325 FTEFSSSP 332
 | : |

RESULTTM 8

Db 268 FREKYDSP 275

US-09-984-130-43

; Sequence 43, Application US/0984130

; Publication No. US20030055231A1

; GENERAL INFORMATION:

; APPLICANT: Ni et al.

; TITLE OF INVENTION: 12 Human Secreted Proteins

; FILE REFERENCE: PF489P2

; CURRENT APPLICATION NUMBER: US/09/984,130

; CURRENT FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/243,792

; PRIOR FILING DATE: 2000-10-30

; PRIOR APPLICATION NUMBER: 09/836,353

; PRIOR FILING DATE: 2001-04-18

; PRIOR APPLICATION NUMBER: 09/836,353

; PRIOR FILING DATE: 2001-10-29

; PRIOR APPLICATION NUMBER: 60/243,792

; PRIOR FILING DATE: 2000-10-30

; PRIOR APPLICATION NUMBER: 09/836,353

; PRIOR FILING DATE: 2001-04-19

; PRIOR APPLICATION NUMBER: PCT/US99/25031

; PRIOR FILING DATE: 1999-10-28

; PRIOR FILING DATE: 1998-10-28

; PRIOR FILING DATE: 1998-10-28

; NUMBER OF SEQ ID NOS: 149

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO: 43

; LENGTH: 1034

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-984-130-43

Query Match 4 2%; Score 102; DB 9; Length 1189;

Best Local Similarity 20.8%; Pred. No. 7.9;

Matches 92; Conservative 56; Mismatches 156; Indels 138; Gaps 22;

Query 71 QILRKGGIKDENIIVMFDDIAFSENPRPGVILNKPGDSDVYK --GVPKDVTKEAVNV 127

Db 252 EAFOGGRGAKVIMIVTD-GESHSDPDLKEVHQQSERDVNTTRYAVAVLGYYNRRG1NP 310

Query 128 QNFYVNLGNESGYTGGNGKVKSGPNNDNIFTIYADHGAGPLIAMPTGDEVMAKDFNEVL 187

Db 311 ETFNEI--KYLASPDPPDHFFNTDEALKDVTDALGRIFs-----L 352

Query 71 QILRKGGIKDENIIVMFDDIAFSENPRPGVILNKPGDSDVYK --GVPKDVTKEAVNV 127

Db 252 EAFOGGRGAKVIMIVTD-GESHSDPDLKEVHQQSERDVNTTRYAVAVLGYYNRRG1NP 310

Query 128 QNFYVNLGNESGYTGGNGKVKSGPNNDNIFTIYADHGAGPLIAMPTGDEVMAKDFNEVL 187

Db 311 ETFNEI--KYLASPDPPDHFFNTDEALKDVTDALGRIFs-----L 352

Query 188 EKMHRKRKYNKVIYVEACESG ---SMFEGLIKKNLNITYATAANKESSKGWVYCP --E 241

Db 353 EGTNK---NETSFLEMSONTGSSHVVEDGVLLGAVGAYDWNGAVLKTSAGKVIPRE 408

Query 188 EKMHRKRKYNKVIYVEACESG ---SMFEGLIKKNLNITYATAANKESSKGWVYCP --E 241

Db 353 EGTNK---NETSFLEMSONTGSSHVVEDGVLLGAVGAYDWNGAVLKTSAGKVIPRE 408

Query 242 SY ---PPPSEIGCLGDTFSISWLEDSDLHDMSKETLEQQHVVKRVG ---SDVPE 293

Db 409 SYLKEFPEELKHAGYLGTVT-----

Query 242 SY ---PPPSEIGCLGDTFSISWLEDSDLHDMSKETLEQQHVVKRVG ---SDVPE 293

Db 409 SYLKEFPEELKHAGYLGTVT-----

Query 294 TSH ---VCREFGT-----

Db 448 FRHFGKVILETMHNRSLTHQHQRGQQIGSYFG-----

Query 294 TSH ---VCREFGT-----

Db 448 FRHFGKVILETMHNRSLTHQHQRGQQIGSYFG-----

Query 342 RDIPLLYLRQKIQKAPMGLESKEAQK -----

Db 496 -DVLLV-----

Query 342 RDIPLLYLRQKIQKAPMGLESKEAQK -----

Db 496 -DVLLV-----

Query 386 DILRUSVKQINVNLLTSTRTGQLVDWDCCFKTLVNSFKNHCGAT-VHYGLKTYGALA 444

Db 548 SVRDLNQDSYNDV-----

Query 386 DILRUSVKQINVNLLTSTRTGQLVDWDCCFKTLVNSFKNHCGAT-VHYGLKTYGALA 444

Db 548 SVRDLNQDSYNDV-----

Query 445 NICNMGVYDVKOTVSAIEQACSM 466

Db 586 K-----TPQRITASELATGL 601

Query 445 NICNMGVYDVKOTVSAIEQACSM 466

Db 586 K-----TPQRITASELATGL 601

RESULTTM 10

US-09-991-258-3

; Sequence 3, Application US/09991258

; Patent No. US2014195A1

; GENERAL INFORMATION:

; APPLICANT: Olmsted, Robert

; APPLICANT: Keith, Paula

; APPLICANT: Dryga, Sergey

; APPLICANT: Caley, Ian

; APPLICANT: Maughan, Maureen

; APPLICANT: Johnston, Robert

; APPLICANT: Davis, Nancy

; APPLICANT: Swanson, Ronald

RESULTTM 9

US-09-984-130-35

; Sequence 35, Application US/0984130

; Publication No. US20030055231A1

; GENERAL INFORMATION:

TITLE OF INVENTION: ALPHAVIRUS VECTORS AND VIROSOMES WITH MODIFIED HIV GENES FOR USE
 :
 TITLE OF INVENTION: VACCINES
 :
 FILE REFERENCE: 01113_0001U3
 :
 CURRENT APPLICATION NUMBER: US/09/991,258
 :
 CURRENT FILING DATE: 2001-11-16
 :
 PRIORITY APPLICATION NUMBER: 09/902,537
 :
 PRIOR FILING DATE: 2001-07-09
 :
 PRIORITY APPLICATION NUMBER: 60/216,995
 :
 PRIOR FILING DATE: 2000-07-07
 :
 NUMBER OF SEQ ID NOS: 19
 :
 SOFTWARE: FastSEQ for Windows Version 4.0
 :
 SEQ ID NO 3
 :
 LENGTH: 2492
 :
 TYPE: PRT
 :
 ORGANISM: Artificial Sequence
 :
 FEATURE:
 :
 OTHER INFORMATION: Description of Artificial Sequence: No. US20020141975A1
 :
 OTHER INFORMATION: synthetic construct
 :
 US-09-991-258-3

	Query Match	Score	Length
Best Local Similarity	4.1%	100	5
Local Similarity	19.7%	10	2492;
Matches	59	31	; Pairs
Matches	77	Mismatches	119; Gaps 18;
Qy	17	LILFSABSRKTOLLNDVESSDKSAGKTPWAVLAVGSNEYNNR-HOADICHAYQLRK 75	
Db	1293	LFLFGYDKRKTHNPKYLSS------LILNTYGSRLHEAGCAPSYHVR- 1337	
Qy	76	GGLKDENITIVMDDAFTSENPRGVIIN-----KPDG---EDVYKGVPKDYTKBAVN 126	
Db	1338	-----GDATATE----GVINAANSKGQPGGGVGALKKFPESEFDLQPIE 1380	
Qy	127	VQNFNYVLNLGENGSVNGNGKVKVS-GPDNDNIFIYVADHGAPGLIAMPGEVMAKFNE 185	
Db	1381	V-----GKARLVAQAKHIIHAGNPNN-----KVSEVEGDSQLAEX-E 1419	
Qy	186	VLEKMRHKKINKMVYVACEGSMEFG--ILKKNLNITYATAANKESSMGVYPE- 241	
Db	1420	SLAKIVNDNNYKSVAI--PLSTGKPSGNKDRULTSQN-HLLTALDTDAVAYCRDK 1475	
Qy	242	-----SPPPPSEI-----SPPPPSEI-----SPPPPSEI-----SPPPPSEI-----	
Db	1476	KWEMTMLKEAVARREAVEEYCISDDSSVTEDAELVRHPKSSLAGRKCYSTSDGKTSYLV 1535	
Qy	264	EDSDLHMKSKELEQQHYVKRRVGSVPETSHWCREFEGTEKMLKDLYSSYIGRNPENDNF 323	
Db	1516	EGTKFHQAOKDIAEIN--AMWPATEANB-QVUMY---ILGEAMSIRSKCPVSESE 1586	
Qy	324	TETTESFSSPISNSGLYVNPRDPLYLQRKIQ 354	
Db	1587	ASTPPSTLPCLCIHAMTPVERVORKASRPEQ 1617	

RESULT 11

	Sequence 4	Application US/09816685
Patent No.	US20020053091A1	
GENERAL INFORMATION:		
APPLICANT: CHANDRAMOULISWARAN, Ishwar et al		
TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR		
FILE REFERENCE: CLO00856		
CURRENT APPLICATION NUMBER: US/09/816,685		
CURRENT FILING DATE: 2001-03-26		
NUMBER OF SEQ ID NOS: 4		
SOFTWARE: FastSEQ for Windows Version 4.0		
SEQ ID NO 4		
LENGTH: 828		
TYPE: PRT		
ORGANISM: Carassius auratus		
US-09-816-685-4		

RESULT 13
 US 09-764-870-313
 ; Sequence 313, Application US/09764870
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PTZ14
 ; CURRENT APPLICATION NUMBER: US/09/764,870
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - consult PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 646
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 313
 ; LENGTH: 707
 ; TYPE: PRP
 ; ORGANISM: Homo sapiens
 US-09-764-870-313

Query Match 4.0%; Score 99; DB 10; Length 707;
 Best Local Similarity 20.6%; Pred. No. 6.7; Mismatches 57; Indels 138; Gaps 22;

Qy 71 QIRKGLKDENTIVWDDIAFSENPRPGYLINKPDGEDYK--GVPKDYTKEAVNV 127
 Db 268 EAQKGGRGAKRVMIVTD-GESHSDPDLKERYTQQSERDNYTRAYAVLGTYNRRGGINP 326
 Qy 128 QNQYVNLGNESVTGGGPNDNITYADIGAPGLAMPIDEMAKDFNEVL 187
 .327 ETPLNEI-----KYISDPDDKKHFFNTDEAALDIVDALGDRFS-----L 368
 Qy 188 EKHKKRKYNNKMYVIYEACESG---SMFEGTIKKNNIYAYTAANSKESSWGVYCP-E 241
 Db 369 EGINK---NEYSGLEMSQTGSSHVVEDGVILGAYDNGAVIKETSAKVIPLRE 424
 Qy 242 SY----PPPSSEITCLGDTFSISWLIEDSLHDMSKTELQQYHVVKRRVG---SDVPE 293
 Db 425 SYLKEFPELKHNHGAYLGTVT-----SVMSRSROGVYVAGAPR 463
 Qy 294 TSH---VCRGFT-----EMMLKDULSSYIGRNPDENDFTTESSSPSSNSGLVNP 341
 Db 464 FNHTGKVLTFTMHNRSLTIHQAMRGQOIGSFIG-----SEITSVDIDGDTY-- 511
 Qy 342 RDPLPLLYLQRKTOKAPMGSLSEKAQK-----KLLDEKHNHRK-QIDQSIT 385
 Db 512 -DILV-----GAPMFNEGGERGVYVYERQNLFVINGTLKOSHYSQNARFGSSIA 563
 Qy 386 DILRLSYKOTVNVLNLSTRTGQPLYDWDCCFKTLVNSFKHNGAT-VHYGLKTYGALA 444
 Db 564 SVRDLNQDSYND-----VVGAPLED-----NHAGATYIFHG--FRGSIL 601

RESULT 14
 US 08-945-749-3
 ; Sequence 3, Application US/08945749
 ; Patent No. US20020136880A1
 ; GENERAL INFORMATION:

; APPLICANT: GARDNER, Richard C
 ; APPLICANT: MACDIARMID, Colin W
 ; APPLICANT: HAY, Robert J
 ; APPLICANT: Auckland Uniservices Limited
 ; APPLICANT: New Zealand and Pastoral Agriculture Research Institut
 ; TITLE OF INVENTION: ALUMINUM RESISTANCE GENE
 ; FILE REFERENCE: 08/945.749
 ; CURRENT APPLICATION NUMBER: US/08/945,749
 ; CURRENT FILING DATE: 1998-01-12
 ; EARLIER APPLICATION NUMBER: PCT/NZ96/0035
 ; EARLIER FILING DATE: 1996-05-01
 ; EARLIER APPLICATION NUMBER: NZ 272039
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 3
 ; LENGTH: 969
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 US-08-945-749-3

Query Match 4.0%; Score 98.5; DB 8; Length 969;
 Best Local Similarity 18.6%; Pred. No. 12; Mismatches 70; Indels 143; Gaps 19;
 Matches 87; Conservative 70; Gaps 19;

Qy 30 LNDNDV-----ESSDKSAGKGTRAVVLVAGSN---EYNNYRHQA---DIC 67
 Db 267 VNDNSKRKKRKGSSDDNSNNKNSKSTS---SDSNDDEDEYNSRPSSSSLSSNNSSLDVC 261
 Qy 68 -----HAYQILRKKGGLKDENIVFMYDDIAFSENPRPGV 102
 Db 262 LYLDDEGSEVPKAWPDCTVLEFSKEETERLSQAIQDAEAHFQYDEDEEDGTSNEDGI 321
 Query Match 4.0%; Score 98.5; DB 8; Length 969;
 Best Local Similarity 18.6%; Pred. No. 12; Mismatches 70; Indels 143; Gaps 19;
 Matches 87; Conservative 70; Gaps 19;

Qy 103 TINPKDGEDVYKGPVKDYTHEAVNQNFN-----VLGNESGVYGN 145
 Db 322 LFSKPIVTN-----DVPELGNRVRVNETENLKGRLRPKRIAPWHLIQRPMLVGLGSNS-TKDS 377
 Qy 146 GKVKSGPNDNIF----IYAH-----GAPGLIAMPT----- 174
 Db 378 KSIQSGLQDNLUVGRNQYPHTISNNPEHFRTYPRVLDSTVHSPTISGLQPQKF 437
 Qy 175 GDEVMAKDFNEVLEKMHKKKKYNNKMYVYACEGSMFEGILKKNNIYAVTRANSKES 234
 Db 438 QDLFVASIYSODNSAGHKTHNNSPTEGKA-ETVSQOLGTAKN-----PSTLSS 487
 Qy 235 WGQYCPESYPP-----PSEIGTCJGDTFSISWLEDSL-----HDMSKTELQYHVK 284
 Db 488 MSYANIEDPPFWLDSNPTEEEMKILSKAFGTHPLTTEDIFGEVREKVELFDYLLIC 547
 Qy 285 RRVGSDVPTESHVCREGTEKMLKDLYSSYIGRNPDENDFTTESSSPSSNSCLVNPRD- 343
 Db 548 FR-SEDIVAKHHRRRRKEQQSATL-----DHEISLSRKQAYGMNSNANNNS 600
 Qy 344 -----TPLLQLRKIQKAPMGSLSEKAQKLLDEKHNHRKOIDOS 383
 Db 601 TSNASRSKWLPSLARRRSSANRNTNTSSSKYRRV-KSEKKMEN 647

RESULT 15
 US-09-759-508B-2
 ; Sequence 2, Application US/09759508B
 ; Publication No. US20020182599A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fishman, Mark C
 ; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease
 ; FILE REFERENCE: 00786/381002
 ; CURRENT APPLICATION NUMBER: US/09/759,508B
 ; CURRENT FILING DATE: 2001-01-12
 ; PRIOR APPLICATION NUMBER: US 60/175,787
 ; PRIORITY FILING DATE: 2000-01-12
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO: 2

```

; LENGTH: 26926
; TYPE: PRT Homo sapiens
; ORGANISM: Homo sapiens
us-09-759-508B-2

Query Match 4.0%; Score 98.5%; DB 9; Length 26926;
Best Local Similarity 19.0%; Pred. No. 1.5e+0;
Matches 101; Conservative 81; Mismatches 152; Indels 197; Gaps 28;
Db 19621 YTYEYKKS--DTDWATSIQSLRGTYTISLTLGAEYVERVKS-----*VNKGAS 19669
Qy 80 D-----ENIIVFMD-----IAFSENPRPGVLINKPDG 109
Db 19670 DPSSSDPFIQAKAREEEPLFDI0SEMRLTLYIKAGASFTMTPFPG-RPVPNVWLSKPD 19728
Qy 110 -----EDVKGVPKDIXTEKAVNQNCFYINVNLIGNESGVTGGNGKVK 150
Db 19729 DLTRRAYVDTDSRTSLTIEANRNDGKTYLTIONVLSAASLT-----VVK 19776
Qy 151 ---SGPDNIFYYA-----DGAPGJAMPTGDEVMADKDFNEYJEKHK 192
Db 19777 VLDTPGPPTNTIVQDVTKEASAVLSDWDPENDGAP-----VKNYH-TEKREA 19822
Qy 193 RKK-----YNIKMYIYEACEBSGSNFEGILKRNUNIYTAAANS-----KESWGVY 238
Db 19823 SKRAWSVTNNCNRSLSYKVNUQEGAY-----YFRVSGENEFVGSGIPATEKEGVK 19873
Qy 239 CPSSYPPPPSETG---TCLGDTFSISWLEDSLHDMSKETLEQQYHVVKRVEGSDVPEPSH 296
Db 19874 ITRK-PSPEKLGTTSKDSVSLTWKPE-HDGSSRIV--HYVE-----19915
Qy 297 VCRFGTERMLKDYLSSTYGRNP-----ENDNFTPTESFSSPISNSGLVNPDIPLY 348
Db 19916 ---ALEBGQKWNWKCAVAKSTHVVSGLRNSEY_FERVAE--NOAGLSDPBLILLPV 19968
Qy 349 LQRKIQKAPMGSLLESKEAQKLLDEKNRHQKIDQDOSTIDLRSVKOTNVNLJSTRTG 408
Db 19969 LKEB-----QLEPPE----IDMRNF----PSHTVYVRAG-----SNLKVDIPISG 20005
Qy 409 QPVVDDNPFCKTLYVNSEKHNHCATVYGLKTGALANICNGVYDVKOTVSA 459
Db 20006 KPLP----KVTLSRDGVLKATMRFNTEITAE---NLINLRESVTA 20045

```

Search completed: May 27, 2003, 15:36:36
 Job time : 76 secs

Db 400 GOLFGIEKGTELLDVRPAGSPPLVDNWLCKTIVKTFETHCGSLSQYGMKHMRSFANIC 459
 Qy 448 NMGYDVKQVYSAEQACCS 465
 Db 460 NAGIPNEPMEEASQAACA 477

RESULT 2
 S51117 Cysteine proteinase (EC 3.4.22.) precursor [similarity] - sweet orange
 C;Species: Citrus sinensis (sweet orange)
 C;Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 09-Jun-2000
 A;Accession: S51117
 R;Alonso, J.; Granell, A.
 submitted to the EMBL Data Library, January 1995
 A;Description: Cloning and expression of an ethylene-related cDNA from orange flavedo de
 A;Reference number: S51117
 A;Status: preliminary
 A;Molecule type: mRNA
 A;Residues: 1-494 <ALP>
 A;Cross-references: EMBL:247793; PID:9633184;
 C;Superfamily: legumain
 C;Keywords: cysteine proteinase; glycoprotein; hydrolase
 F;1-22/Domain: signal sequence #status predicted <SIG>
 F;23-55/Domain: propeptide #status predicted <PRO>
 F;56-494/Product: cysteine proteinase #status predicted <MAT>
 F;151,336/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 51.4%; Score 1262.5; DB 2; Length 494;
 Best Local Similarity 51.6%; Pred. No. 1..5e-79;
 Matches 238; Conservative 69; Mismatches 120; Indels 9; Gaps 3;

Oy 30 LNDDVESSDKSAGTRAVLYVGSNEVYNNQHADICHAYAOLLRKGGLKDENTIVFMDY 89
 Db 50 DWD----DYGTRAVYLGLGSNOFWNTRHQDITCHAOLLRKGGLKDENTIVFMDY 103
 Qy 90 DIAFSSENPRPGTINKPDKDYEYKGVPDKDTKEAVRNQVYNVLGNGESCVTGNGKVV 149
 Db 104 DIAFNEENPRPGTINHPHDWYKGVPDKDGTEDVTEKFVAVLGNKPAALTGGSKVV 163
 Qy 150 KSGPNDNTIYYADHAGPGLIAMPDTGEVMAKDFNEVLEKHKRKKYKVNKAWIYVEACESG 209
 Db 164 DSGNDNDHIFTFYSDHGEGGVGLPMTTSRYIAYEALIDLVKKHASNQYKSLVYPLACESG 223
 Qy 210 SMFEGILKNNINYAVTRANSKESSWGVYCPESSYPPSEITGTCGDTFTSISWLEDSDLH 269
 Db 224 SIFGGLLGLNLNYATTASNAEBSWNGTYCPCBGPPEYSTCLGDLYSIAMWEDSDIH 283
 Qy 270 DMSELTQHYYHVKRGSIDYFETSHVCRFOTEMKLDYLISSYIGRNPDENDNFFTEF 329
 Db 284 NRTPETLHQYEVLKTRASYSYGHSMQTDIGUSKRNLTYLGTNPANDNTYEVDEN 343
 Qy 330 S-SPIINSCLVNDRDIPILYLORKIQKAPMGLESKEAQKLLDEKNHRKOIDOSSTDIL 388
 Db 344 SLRPASKA-VNGRADDLHFNWDKYRAPEPTPKRAQEAOFFEEAMSHRHVDRHTRKLIG 401
 Qy 389 RLSVKQTQTNVLNLDSTRGQPLVWDWCFTLVLNSFKNHGATVHYGLKTYGALANIC 448
 Db 402 KLFGLIEKGPELTNTVREAGQPLVDDGKURSLVRLTFFESHGALSQYGMHRMSLANIC 461
 Qy 449 MGVDVKQTVPSAIQAC 464
 Db 462 TGICKERMAEASQAAC 477

RESULT 3
 T05302 vacuolar processing enzyme (EC 3.4.22.-) isozyme gamma precursor - Arabidopsis thaliana
 N;Alternative names: protein F26P21.60
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jun-2000
 C;Accession: T05302

R;Bevan, M.; Terryn, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.;
 ewes, H.W.; Mayer, K.F.X.; Schueler, C.; Submitted to the Protein Sequence Database, October 1998
 A;Reference number: Z15407
 A;Accession: T05302
 A;Molecule type: DNA
 A;Residues: 1-494 <BEV>
 A;Cross-references: EMBL:AL031804
 A;Experimental source: cultivar Columbia; BAC clone F26P21
 C;Genetics:
 C;Map position: 4
 A;Introns: P26P21.60
 A;Note: F26P21.60
 C;Superfamily: legumain
 C;Keywords: cysteine proteinase; glycoprotein; hydrolase
 F;1-27/Domain: signal sequence #status predicted <SIG>
 F;28-53/Domain: propeptide #status predicted <PRO>
 F;54-494/Product: vacuolar processing enzyme isozyme gamma #status predicted <MAT>
 F;336/Binding site: carbohydrate (Asn) (covalent) #status predicted
 . Query Match 50.8%; Score 1246; DB 2; Length 494;
 . Best Local Similarity 53.6%; Pred. No. 2.1e-78;
 . Matches 233; Conservative 71; Mismatches 121; Indels 10; Gaps 4;
 . Qy 36 ESSDKSAKGTRAVLYVAGSNEVYNNQHADICHAYQILRKGLKDENTIVFMDYAFSS 95
 Db 49 ENDDDSNSCTRWAVLASSGYNWYQKRDICHAYQILRKGLKEEENIVVPMYDDANNY 108
 Qy 96 ENPRRGVILNKPKDCEYKVPKDYTKEAQNQVNFYVLLGNECVTGNGKVKSGPND 155
 Db 109 ENPRGTINSPHCKDVQGPDKTGYGDVNDLNFLAVILGDKTAVYGGSGRVVDGPND 168
 Qy 156 NIFTYYADHAGPGLIAMPDTGEVMAKDFNEVLEKHKRKKYKVNKAWIYVEACESG 215
 Db 169 HIFIFYSDHGEGPVGLMPMTSPYLYANDLNVLKHKHALGTKSLSVLFLEAESGSIFEGL 228
 Qy 216 LKKNLNTYAVTAANSKESSWGVYCPESYPPIPSEITGLGDTFSISWLEDSDLHOMSKET 275
 Db 229 LPESUNIYTATSNAAEESSWGTVCPGBPSPPEYEVGLDOLYSVAMMEDSGMHNQLET 288
 Qy 276 LEQQYHVYKRR--VGSDDVPETSHVCRGTEKMLKDLYSSYIGRNPDENDNFFTE-SFS 330
 Db 289 LHQQYELVKRTAPVGYY-GSHVMQYGVICSKDNLDLMTGNTNANDNTFADANSLK 346
 Qy 331 SPISNSGLVNPDRDIPLYLQRKLOKAPMGSLSESKEAQKLLDEKNHRKOIDOSSTDILRL 390
 Db 347 PP--SRVYNQDADLWVKEYRKAPGSAKTEAKQVDEAMSIRLHINSVLVGKI 403
 Qy 391 SVKOTNVNLILSTRTGQPLVWDWCFTLVLNSFKNHGATVHYGLKTYGALANIC 450
 Db 404 LFGTISRGPPLEVNLKVRSAGQPLVDDWNLKVNRVAFERHGSSLQXQIKHMRSFANICNAG 463
 Qy 451 VDKQTVSAEQAC 465
 Db 464 IQHEOMEAEASQACT 478

RESULT 4
 T12043 probable legumain (EC 3.4.22.34) precursor - kidney bean endopeptidase; bean endopeptidase; bean endopeptidase (kidney bean)
 C;Species: Phaseolus vulgaris (kidney bean)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 21-Jul-2000
 R;Senyuk, V.; Rotari, V.; Becker, C.; Zaharov, A.; Muentz, K.; Vaintraub, I.; Horstma
 Eur. J. Biochem. 255, 546-558, 1998
 A;Title: Does an asparagine-specific cysteine endopeptidase trigger phaseolin degrad
 A;Reference number: Z17389; PMID:99089618; PMID:9874222
 A;Accession: T12043

A>Status: preliminary; translated from GB/EMBL/DDJB
 A;Molecule type: DNA
 A;Residue: 1-484 <SEIN>
 A;Cross-references: EMBL:Z99956; NID:92511696; PID:92511697

A; Experimental source: cultivar Moldavian; cotyledon; clone cp6b
 C; Function: responsible for the initiation of phaseolin proteolysis

A; Description: responsible for the initiation of phaseolin proteolysis
 C; Superfamily: legumin
 C; Keywords: cysteine proteinase; glycoprotein; hydrolase

F; 1-25/Domain: signal sequence #status predicted <SIG>
 F; 26-44/Domain: propeptide #status predicted <PRO>

F; 45-84/Product: probable legumin #status predicted <MAT>
 F; 326/Binding site: carbohydrate (Asn) (covalent) #status predicted

R; Lin, X.; Kaul, S.; Rounseley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y
 M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umeyam, L.; Venter,
 C.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
 Nature 402, 761-768, 1999
 A; Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.
 A; Reference number: A84420; MUID:20083487; PMID:10617197

A; Accession: FJ4654
 A; Status: preliminary
 A; Molecule type: DNA

A; Residues: 1-478 <STO>
 A; Cross-references: GB:AE002093; NID:93413718; PIDN: AAC31241.1; GSPDB: GN00139

C; Genetics:
 A; Gene: T19L18_25; At2g25940
 A; Map Position: 2
 A; Introns: 61/3; 116/3; 169/1; 197/3; 280/3; 350/3; 418/3

C; Superfamily: legumin
 C; Keywords: cysteine proteinase; glycoprotein; hydrolase
 F; 1-20/Domain: signal sequence #status predicted <SIG>
 F; 21-41/Domain: amino-terminal propeptide #status predicted <PRO>
 F; 42-478/Product: vacuolar processing enzyme isozyme #status predicted <MAT>
 F; 135-321/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match Score 1243.5; DB 2; Length 484;
 Best Local Similarity 51.6%; Pred. No. 3.1e-78;
 Matches 238; Conservative 78; Mismatches 132; Indels 13; Gaps 5;

Qy 10 ILVFLHALLTSEASRKTQLIND----NDVESSDKSAKTRWAVLVAGSNEYNNYRHQA 64
 Db 14 LLFELFLVALV---SAGRDLVGFDFLRPLPSDNGD-NVHGTRWALLFLAGSSGYNNYRHQA 67

Qy 65 DICHAZOILRKGGGLDENITYFMDDIAFSSENPRPGVINKPQGDVEDYKGVPKDYTKEA 124
 Db 68 DICHAZOILRKGGGLDENITYFMDDIAFNSENPRGVINSPNQDDEVYKGVPKDYGED 127

Qy 125 VNVQNFVNVLIGNESGVTGGNGKVKYSGPNDNIFTIYADHGAPGLIAMPQGDEVMAKDFN 184
 Db 128 VTAHNPYAALLGDKSKLTGGSKVWNNGPNDHFIFYSDFIGGPVGSPAGPYIASDLN 187

Qy 185 EVLEKKHKRKVKNWVYVEACESSGMFEGILLKRNLIYAVTAANSKESSNGVYCPESYP 244
 Db 188 EVLKKKHASGTYKNUVYLEACESSGMFEGILPENVYATTASADESSWGTCPGEDP 247

Qy 245 PPSPEIETCLGDTFSISWLSDLSHMSKTEQQYHVYKRR-VGSDVPETSHYCREFGTE 303
 Db 248 SPPPEYSTECLGDTLSDLYSAWMDSDRHLRTELHQKQKLKVERTISGGLYGSHVQYQDV 307

Qy 304 KMLKDYLSSYIGRNPNENDNFTTESFSPPSISNSGLYNPRDIPLYLQRKIQKAPNGSLES 363
 Db 308 GLSKDLEHYLGTDPANENUTFDE-NSLNSSKAVNORDAHLWFDKERKAPGSPKK 366

Qy 364 KEAQKLLDEKNRKHQDOSITDILRLSVKOTNVNLSTRTSQPLVDDWDCFKTLVN 423
 Db 367 NEARKOYLEVMSRHMHDSSVELVGRLFGTEKAPBLNAVRPAGSALVDDWCLKTMVR 426

Qy 424 SFKNRGCGATHYGLKTGALANICNNMGVDQTVSAEQAQC 464
 Db 427 TFEETHCGSLISQYGMKMRSEFNMCNVGNIKEQMREASAQC 467

RESULT 5
 T02629 vacuolar processing enzyme (EC 3.4.22.-) isozyme alpha precursor - *Arabidopsis thaliana*

N; Alternate names: protein T19L18_25
 C; Species: *Arabidopsis thaliana* (mouse-ear cress)
 C; Date: 24-Mar-1999 #sequence_change 24-Mar-1999 #text_change 16-Feb-2001

C; Accession: T02629; S60049; F84654
 R; Rounseley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, submitted to the EMBL Data Library, August 1998

A; Description: *Arabidopsis thaliana* chromosome II BAC T19L18 genomic sequence.
 A; Reference number: 214681
 A; Accession: T02629

A; Status: translated from GB/EMBL/DDBJ
 A; Molecule type: DNA
 A; Residues: 1-478 <ROU>
 A; Cross-references: EMBL:AC004747; NID:93413696; PID:93413718

A; Experimental source: cultivar Columbia
 R; Kinoshita, T.; Nishimura, M.; Hara-Nishimura, I.
 Plant Mol. Biol. 29, 81-89, 1995
 A; Title: Homologues of a vacuolar processing enzyme that are expressed in different org-

A; Reference number: S60049; MUID:96017615; PMID:7579169
 A; Accession: S60049
 A; Molecule type: DNA
 A; Residues: 1-69, 'R', 71-83, 'L', 85-90, 'E', 92-478 <KIN>
 A; Cross-references: EMBL:D61393; NID:9110444; PIDN:BA09614.1
 A; Note: the authors translated the codon ATT for residue 112 as Tyr and the codon ATT for

RESULT 6
 S60050 vacuolar processing enzyme (EC 3.4.22.-) isozyme beta precursor - *Arabidopsis thaliana*
 C; Species: *Arabidopsis thaliana* (mouse-ear cress)
 C; Date: 06-Dec-1996 #sequence_change 06-Dec-1996 #text_change 21-Jul-2000

C; Accession: S60050
 R; Kinoshita, T.; Nishimura, M.; Hara-Nishimura, I.
 Plant Mol. Biol. 29, 81-89, 1995
 A; Title: Homologues of a vacuolar processing enzyme that are expressed in different o-

A; Reference number: S60049; MUID:96017615; PMID:579169
 A; Accession: S60050
 A; Molecule type: DNA
 A; Residues: 1-184 <KIN>
 A; Cross-references: EMBL:D61394; NID:9110446; PIDN: BAA09615.1; PMID:91805364

A: Note: the authors translated the codon TAC for residue 135 as Leu and the codon TAC fd

C: Genetics: 68/3; 123/3; 176/1; 204/3; 271/2; 287/3; 355/3; 422/3

C: Keywords: cysteine proteinase; glycoprotein; hydrolase

C: Superfamily: legumain

F: 2-1/Domain: signal sequence #status predicted <SIC>

F: 4-2-46/Domain: amino-terminal propeptide #status predicted <PRO>

F: 362-484/Domain: carbohydrate (Asn) (covalent) #status predicted

F: 307/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 46.8%; Score 1148; DB 2; Length 484;

Best Local Similarity 48.7%; Pred. No. 1..2e-71;

Matches 228; Conservative 69; Mismatches 153; Indels 18; Gaps 5;

Query 7 HFQILYFLHAILFSAESR-----KQQLANDNDVESSDSAKGTRWAVLVAGSNYE 57

Db 6 YFRPALLLVLVHAESRGFEPKLMPTEANPAD---QDEOOGVGRWAVLVAGSSGY 62

Query 58 YNRHQADICHAYQILRKGGKLDENIVPKDDIAFSSENPRPGVILINKPDGEDVYKCP 117

Db 63 GNYRHQADVCHAYQILRKGGKKEENIVMLYDDIANHPLNRPSTLHNHPDGDVYAVP 122

Query 64 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFTYYADHGAPGLTAMPGE 177

Db 123 KDTYGTSSVTAANFYAVLQDGOKAVGGSGKVIAKSPNDHIEVYADHGPGVLMGPNTPH 182

Query 178 VMAKDENEVLEKMRKKYKRNMYVITVEACESGSFEGTIGLKUNIYATAANRESSGV 237

Db 183 IYADEFTELUKKKHASGTYKEMVYVEACESGSFEGTIGMPKDUNIYVTTASNAQESSYGT 242

Query 238 YCPESYPSPPESEITGCLGDFTSISKMLESDLHDNSKETLEQYHVKKRPGSDVPETHV 297

Db 243 YCPGMNPSPSEYITOLGDYLSVAMMEDSTHNLKETKQQTHTMRSNTYNSGGSHV 302

Query 298 CRFGTERKMLKSYLSSYIGRNPDENFTESSPISNSGLVNPDRDILYLQRKIQKAP 357

Db 303 MEGNNNSIKSEKFLYLGFPETAVNLPLNLPAK - SKGVVRDADLFLWMMHRPSE 360

Query 358 NGSLSEKQAQKLDENKRKQDIDTDLRLSVKQT NVLNLIUSTRTGTCOPLYWDW 416

Db 361 DGSRKDDTLKELTEPTRHKHLASVELIATILEFGPTMVNLNV -- REGPLPVDDWE 417

Query 417 CFPKLVNSFKHCGATVHYGLKYTGALANTCNGMVDVKOTVSAEQAC 464

Db 418 CLKSMYRVFEHCGSLTQYGMKHMRAFANVCNNVSKELMEEASTAAC 465

RESULT 8

T12044 probable legumain (EC 3.4.22.34) precursor - kidney bean

N: Alternative names: asparagine endopeptidase; bean endopeptidase; phaseolin; vicilin

C: Species: *Phaseolus vulgaris* (kidney bean)

C: Cross-references: EMBL:299957

C: Date: 02-Mar-2001 #sequence_revise

C: Accession: C96652

C: Description: Isolation of cDNA clone encoding legumain-like proteinase (LLP2) from *Arabidopsis thaliana*

R: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ansari, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A: Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.N.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzali, R.; Izzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venturini, J.C.; Davis, R.W.

A: Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A: Reference number: A86141; MUID:21016719; PMID:11130712

A: Accession: C96652

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-536 <SIC>

A: Cross-references: GB:AEE005173; NID:96630462; PIDN:AAF19550.1; GSPDB:GN00141

C: Genetics: F23N19.7 [Imported] - *Arabidopsis thaliana* (mouse-ear cress)

C: Date: 02-Mar-2001 #sequence_revise

C: Accession: C96652

C: Description: Isolation of cDNA clone encoding legumain-like proteinase (LLP2) from *Arabidopsis thaliana* (mouse-ear cress)

R: Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chinn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Ansari, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A: Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; C.N.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marzali, R.; Izzo, M.; Rooney, T.; Rowley, D.; Sakano, H.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venturini, J.C.; Davis, R.W.

A: Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A: Reference number: A86141; MUID:21016719; PMID:11130712

A: Accession: C96652

A: Status: preliminary

A: Molecule type: DNA

A: Residues: 1-536 <SIC>

A: Cross-references: GB:AEE005173; NID:96630462; PIDN:AAF19550.1; GSPDB:GN00141

C: Genetics: F23N19.7

A: Gene: F23N19.7

A: Map position: 1

C: Superfamily: legumain

Query Match 45.8%; Score 1125; DB 2; Length 536;

Best Local Similarity 44.5%; Pred. No. 5.5e-70;

Matches 232; Conservative 68; Mismatches 149; Indels 72; Gaps 8;

Query 7 HFQILYFLHAILFSAESR-----KQQLANDNDVESSDSAKGTRWAVLVAGSNYE 57

Db 6 YFRPALLLVLVHAESRGFEPKLMPTEANPAD---QDEOOGVGRWAVLVAGSSGY 62

Query 58 YNRHQADICHAYQILRKGGKLDENIVPKDDIAFSSENPRPGVILINKPDGEDVYKCP 117

Db 63 GNYRHQADVCHAYQILRKGGKKEENIVMLYDDIANHPLNRPSTLHNHPDGDVYAVP 122

Query 118 K-----DTYKEAVN 127

Db 123 KALHNNSDSDCRDICYGKPIMCGPFIGIAPRFLIATCSVTYVLYKFLFDYDGTGSSVTA 182

Db 124 KKKKHASCTYKEMVIVYEAESGSFEGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 128 QNFYNNVLLGNESGTGGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 129 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 130 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 131 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 132 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 133 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 134 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 135 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 136 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 137 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 138 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 139 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 140 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 141 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 142 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 143 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 144 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 145 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 146 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 147 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 148 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 149 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 150 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 151 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 152 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 153 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 154 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 155 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 156 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 157 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 158 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 159 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 160 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 161 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 162 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 163 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 164 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 165 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 166 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 167 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 168 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 169 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 170 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 171 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 172 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 173 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 174 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 175 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 176 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 177 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 178 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 179 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 180 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 181 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 182 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 183 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 184 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 185 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 186 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 187 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 188 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 189 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 190 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 191 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 192 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 193 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 194 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 195 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 196 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 197 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 198 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 199 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 200 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 201 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 202 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 203 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 204 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 205 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 206 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 207 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 208 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 209 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 210 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 211 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 212 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 213 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 214 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 215 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 216 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 217 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 218 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 219 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 220 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 221 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 222 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 223 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 224 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 225 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 226 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 227 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 228 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 229 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 230 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 231 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 232 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 233 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 234 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 235 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 236 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 237 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 238 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 239 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 240 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 241 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 242 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 243 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 244 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 245 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 246 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 247 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 248 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 249 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 250 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 251 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 252 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 253 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 254 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 255 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 256 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 257 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 258 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 259 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 260 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 261 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 262 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 263 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 264 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 265 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 266 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 267 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 268 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 269 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 270 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 271 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 272 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 273 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 274 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 275 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 276 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 277 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 278 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 279 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 280 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 281 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 282 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 283 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 284 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 285 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 286 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 287 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 288 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 289 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 290 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 291 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 292 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 293 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 294 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 295 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 296 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 297 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Db 298 KDTYKEAVNYNFVNLLGNESGTGGNGKVKSGPNDNIFYYADHGAPGLIAMPTGDEVIAKDFNEVL 187

Query 299 KDTYKEAVNYNF

Qy	94 SSENPRPGVIIINKPDGEDVYKGPDKYTKEAVNVONFYNVNLLGNESTGGNGKVKVKGSP 153	Qy	215 IURKLNLYIAYTAANSKESWGVYCPESYPPPEIGTCLGDTFSISWLIEDSDLHDMSK 274
Db	106 HEINPRPGVIIINKPDGEDVYKGPDKYTKEAVNVONFYNVNLLGNESTGGNGKVKVKGSP 165	Db	231 IMPKDVIDYVTASNAQESWGVYCPGMESPSPPEFTICLGDTFSISWLIEDSDLHDMSK 290
Qy	154 NDNIFIYAADHGPGLIAMPTGDEVMAKDFNEVLEKMHKRKKYNNMVIYVEACESGSMEF 213	Qy	275 TLEQQHYVVKRVRG--SPDPETSHVCRFGTEKMLKDYLSSYIGRNPNENDNFTTESFSSP 332
Db	166 EDTRFVYSDHGGPGVLGMPNMHLYAMDFDWLKKKHASGGYKEMVYVENCEGSFIE 225	Db	291 TVKQYQSSVKARTSNNTYAAAGSHVMQINGNOSIKADKLFLFGFDPSVNF-----P 342
Qy	214 GIURKLNLYIAYTAANSKESWGVYCPESYPPPEIGTCLGDTFSISWLIEDSDLHDMSK 273	Qy	333 ISNSGL-----VNPRDIPPLYLRQKIQAPGSLESKEAQKQLDEKHNHQIDQSTD 386
Db	226 GIMPKDNLNLYIAYTAANSKESWGVYCPESYPPPEIGTCLGDTFSISWLIEDSDLHDMSK 285	Db	343 PNHHLNAPMEVNORDAELHPMWOLYKRSENGSEKKEILQOIQDAKHRSHLDSSMOL 402
Qy	274 ETLEQQHYVVKRVRG--DVPETSHVCRFGTEKMLKDYLSSYIGRNPNENDNFTTESFSS 331	Qy	387 IRLRSVKOTINVNLNLSTRTTGPOLVDDCFKTLVNSFKHICGATHYGLKVTGALAN 446
Db	286 ESEQQYQSYVKQTSNFEAYAMSHVMQGDAMTAENLYLHGFDATPVNE-----337	Db	403 IGDLFLGPKKASAIILKSVERPSPLVDWGCILKSMVRVETCCGSLTQYGMKHMRTFANI 462
Qy	332 PISNSGL-----VNPRDIPPLYLRQKIQAPGSLESKEAQKQLDEKHNHQIDQSDIT 385	Qy	447 CNMGDVYRQTVSAIEQAC 465
Db	338 PPNGRLRKSMEVYNORDAELFMWVOYQRSNHLPERKTDILKQIEEVKHKRKHLDSSVE 397	Db	463 CNAGVSHTSMEERCAAC 481
Qy	386 DIRLSVKQTINVNLNLSTRTTGPOLVDDCFKTLVNSFKHICGATHYGLKVTGALAN 445		RESULT 10
Db	398 LIGVLLYGEPEKASSVLRSVTFQPLVDDWTCLIKSMVRVYETHCGSLTQYGMKHMRAFAN 457		JX0344
Qy	446 ICNNGDVYDQTVSAIEQAC 464		legumain (EC 3.4.22.34) precursor - jack bean
Db	458 ICHSGV---SESMEKAC 472		N;Alternate names: asparaginyl endopeptidase
		C;Species: Canavalia ensiformis (jack bean)	
		C;Accession: JX0344 #sequence_revision 26-May-1995 #text_change 21-Jul-2000	
		C;Accession: JX0344 #sequence_revision 26-May-1995 #text_change 21-Jul-2000	
		R;Takeda, O.; Miura, Y.; Matsushita, H.; Kato, I.; Abe, Y.; Yokosawa, H.;	
		J; Biochem. 116, 541-546, 1994	
		A;Title: Isolation and analysis of cDNA encoding a precursor of Canavalia ensiformis A;Reference number: JX0344; MUID:9515263; PMID:7852272	
		A;Accession: JX0344; Molecule type: mRNA	
		A;Residues: 1-475 <TAK>	
		A;Cross references: DDBJ:D31787; NID:9499293; PIDN:BAA06596..1; PID:9499294	
		A;Experiment source: Seed	
		C;Comment: This enzyme is involved in posttranslational processing of concanavalin A	
		C;Superfamily: lectomain	
		C;Keywords: cysteine proteinase; glycoprotein; hydrolase	
		F;1-15/Domain: signal sequence #status predicted <SIG>	
		F;1-16/Domain: propeptide #status predicted <PRO>	
		F;36-475/Domain: product: legumain #status experimental <MAT>	
		F;300/Binding site: carbohydrate (Asn) (covalent) #status predicted	
		Query Match 44.0%; Score: 1080; DB: 2; Length: 475;	
		Best Local Similarity 48.3%; Pred. No. 5.9e-67;	
		Matches 207; Conservative 73; Mismatches 129; Indels 20; Gaps 4;	
RESULT 9			
Qy	vacuolar processing enzyme (EC 3.4.22.-) precursor - castor bean		
C;Species: Ricinus communis (cassitor bean)			
C;Accession: JX0287			
R;Hara-Nishimura, I.; Takeuchi, Y.; Nishimura, M.			
Plant Cell 5:1651-1659, 1993			
A;Title: Molecular characterization of a vacuolar processing enzyme related to a putatively			
A;Reference number: JQ0387; MUID:8312744			
A;Accession: JQ0387			
A;Molecule type: protein			
A;Residues: 104-141 <H22>			
A;Experimental source: seed			
C;Comment: This enzyme plays a crucial role in the biosynthesis of vacuolar components a			
C;Superfamily: legumain			
C;Keywords: cysteine proteinase			
F;1-31/Domain: signal sequence #status predicted <SIG>			
F;158-497/Domain: propeptide #status predicted <PRO>			
F;158-374/Product: vacuolar processing enzyme #status predicted <MAT1>			
F;158-374/Domain: carboxyl-terminal propeptide #status predicted <CTP>			
F;158/Modified site: pyrollidone carboxylic acid (Gln) (in mature form) #status predicted			
F;320/Binding site: carbohydrate (Asn) (covalent) #status predicted			
Query Match 45.2%; Score: 1109; DB: 2; Length: 497;			
Best Local Similarity 50.1%; Pred. No. 6.3e-69;			
Matches 20; Conservative 62; Mismatches 141; Indels 16; Gaps 3;			
Qy	35 VESSSDKSAKGTRAVLVLAGSNEYNYRHQDICHAYQILRKGLDENIVPMYDDIAFS 94	Qy	224 AVTAANSKESSWGVYCPESYPPPEIGTCLGDTFSISWLIEDSDLHDMSKTELQQHV 283
Db	51 VQVDDDDQGLGTRAVLVLAGSNEYNYRHQDICHAYQILRKGLDENIVPMYDDIAFS 110	Db	218 VTTASNAQENSFGTYPGSHVMQYGDNTAEKLYLHGFDPATVNF-----PHINGNLEAK 277
Qy	95 SENPRPGVIIINKPDGEDVYKGPDKYTKEAVNVNLLGNESTGGNGKVKVKGSP 154	Qy	284 KRRVGDYVET--SHVCRFGTEKMLKDYLSSYIGRNPNENDNFTTESFSPISNSGL-- 338
Db	111 ELNPRLPGVIIINKPDGEDVYKGPDKYTKEAVNVNLLGNESTGGNGKVKVKGSP 170	Db	278 RKRRTNSNSYRPGSHVMQYGDNTAEKLYLHGFDPATVNF-----PHINGNLEAK 329
Qy	155 DNIFTYYADHGPGLIAMPTGDEVMAKDFNEVLEKMHKRKKYNNMVIYVEACESGSMEF 214	Qy	339 ---VNPRDIPPLYLRQKIQAPGSLESKEAQKQLDEKHNHQIDQSDITDILRSVKQT 395
Db	171 DRFLYVYSDHGGPGVGLCPNLPFLYANDFIEVKKHAAGGYKMKVYVEACESGSIFEG 230	Db	330 MEVYNQRDAELFMWQYQRSNHIOPERKTHILEQITETVKHRNHLDSVELGVLYPG 389

Qy 396 NVLNLLSTRITGQPLYDDWDCFKTLVNSFKHCGATVHYGLKYTGALANICNMGVDVVKO 455
Db 390 KSSSVLHSVRAPOLPLVDDWTCILKSMYRVEFHCGSLTQGMKHMRAFNCSGV--- 445
Qy 456 TSVAIEQAC 464
Db 446 SKASMEEC 454

RESULT 11
T07132
cysteine proteinase (EC 3.4.22.-) precursor [similarity] - soybean
C;Species: Glycine max (soybean)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 21-Jul-2000
C;Accession: T07132
R;Shimada, T.; Hiraiwa, N.; Nishimura, M.; Hara-Nishimura, I.
Plant Cell Physiol. 35, 713-718, 1994
A;Title: Vacuolar processing enzyme of soybean that converts proproteins to the corresponding mature proteins
A;Reference number: 215942; MUID:94556350; PMID:8075902
A;Status: preliminary; translated from GB/EMBL/DDJB
A;Molecule type: mRNA
A;Residues: 1-503 <PFS2
A;Cross-references: EMBL:Ad007743
A;Experimental source: cotyledons; clone PHS1
C;Function:
A;Description: Plays a crucial role in the biosynthesis of vacuolar components and responsible for the maturation of proproteins.
C;Superfamily: legumin
C;Keywords: cysteine proteinase; glycoprotein; hydrolase
F:1-18-/Domain: signal sequence #status predicted <SG>
F:39-62-/Domain: propeptide #status predicted <PRO>
F:63-503-/Domain: product; cysteine proteinase #status predicted <MAT>
F:328-/Binding site: carboxylate (Asn) (covalent) #status predicted
Query Match 43.0%; Score 1054; DB 2; Length 503;
Best Local Similarity 45.2%; Pred. No. 4e-65;
Matches 213; Conservative 82; Mismatches 146; Indels 30; Gaps 7;
C;Function:
A;Description: Plays a crucial role in the biosynthesis of vacuolar components and responsible for the maturation of proproteins.
C;Superfamily: legumin
C;Keywords: cysteine proteinase; glycoprotein; hydrolase
F:1-18-/Domain: signal sequence #status predicted <SG>
F:39-62-/Domain: propeptide #status predicted <PRO>
F:63-503-/Domain: product; cysteine proteinase #status predicted <MAT>
F:328-/Binding site: carboxylate (Asn) (covalent) #status predicted
Query Match 43.0%; Score 1054; DB 2; Length 503;
Best Local Similarity 45.2%; Pred. No. 4e-65;
Matches 213; Conservative 82; Mismatches 146; Indels 30; Gaps 7;
C;Function:
A;Description: Plays a crucial role in the biosynthesis of vacuolar components and responsible for the maturation of proproteins.
C;Superfamily: legumin
C;Keywords: cysteine proteinase; glycoprotein; hydrolase
F:1-18-/Domain: signal sequence #status predicted <SG>
F:39-62-/Domain: propeptide #status predicted <PRO>
F:63-503-/Domain: product; cysteine proteinase #status predicted <MAT>
F:328-/Binding site: carboxylate (Asn) (covalent) #status predicted
Query Match 43.0%; Score 1054; DB 2; Length 503;
Best Local Similarity 45.2%; Pred. No. 4e-65;
Matches 213; Conservative 82; Mismatches 146; Indels 30; Gaps 7;
C;Function:
A;Description: Plays a crucial role in the biosynthesis of vacuolar components and responsible for the maturation of proproteins.
C;Superfamily: legumin
C;Keywords: cysteine proteinase; glycoprotein; hydrolase
F:1-18-/Domain: signal sequence #status predicted <SG>
F:39-62-/Domain: propeptide #status predicted <PRO>
F:63-503-/Domain: product; cysteine proteinase #status predicted <MAT>
F:328-/Binding site: carboxylate (Asn) (covalent) #status predicted
Query Match 43.0%; Score 1054; DB 2; Length 503;
Best Local Similarity 45.2%; Pred. No. 4e-65;
Matches 213; Conservative 82; Mismatches 146; Indels 30; Gaps 7;
C;Function:
A;Description: Plays a crucial role in the biosynthesis of vacuolar components and responsible for the maturation of proproteins.
C;Superfamily: legumin
C;Keywords: cysteine proteinase; glycoprotein; hydrolase
F:1-18-/Domain: signal sequence #status predicted <SG>
F:39-62-/Domain: propeptide #status predicted <PRO>
F:63-503-/Domain: product; cysteine proteinase #status predicted <MAT>
F:328-/Binding site: carboxylate (Asn) (covalent) #status predicted
Query Match 43.0%; Score 1054; DB 2; Length 503;
Best Local Similarity 45.2%; Pred. No. 4e-65;
Matches 213; Conservative 82; Mismatches 146; Indels 30; Gaps 7;

Qy 10 ILYFLHALIFLSAESRKTLQ-----LNDNDESSDKSAGCTRWAVLGASNEYNNYRH 62
Db 26 LLDFLSSL--HGSVARPLRWEWPVRLGPEDVADVEDEGTRWAVLGASNGCYNRH 83
Qy 63 QADICHAUYOLRKGGKLKDENITVFMYDIAFSENPRGVITNKPGDEDVYKGVPKDYTK 122
Db 84 QADICHAUYOLRKGGVKEENLVFMYDIAFSENPRGVINHPGPVNIDGVPKDYG 143
Qy 123 EAIVQNQFYNVLLGNECVTGGKVKVSPGDNNIYYADHGAPEGLJAMPTGDEVMAKD 182
Db 144 DFTVADNLAYAVLGDKSJKVRGSGSKVILNSKAEDRIEYSDHGPVGLGMNMPVYAMD 203
Qy 183 FNEVLEKMRKKYKVKHIVYBACESSGMFEGILKLNLYTAVTAANSKESSESWGY-CPE 241
Db 204 FIDVLKKKHAASRYQQVIVYVACESGVFOCIMPDIIVYVTTAASNEERSWGYLVPG 263
Qy 242 SYPPPSEITCLGDTFSISHLDSLHDMSKETLQQHYVVKRRY-GSDVPEPTEHVCR 299
Db 264 VYPASPPETITCUGDLQSVAMWEDSETHNLKETLKQFAVKERTUNNNYGLGSHVTE 323
Qy 300 FGTEKMLRDYLSSYIGRNPDENFTESFSSPSISNGL-----VNPRDIPLYLQRKI 353
Db 324 YGDTNTIDEKLYLHGFDPASYNL-----PPNNGRLESKMEVYNORDAEILFWQMY 375
Qy 354 QKAPMGSLSEKAAQKLLEDEKHNKRQIDQSITDILRSVKQINVNLALTSTRTGQPLVD 413
Db 376 QRLDHQSEKKRDLKLKISETVKRNHLHDGTSVLEGIVLFFGTPRGSSQLQSVRASGLPLVD 435
Qy 414 DWDFKTLVNSFKNHGATVHGLKTGALANICNMGVDMQVTSATEQAC 464
Db 436 DWECLKSRSVRFETHCGSLTQGMKHMRAFNCRG-----SEDLMEETC 482
Qy 336 SGL-----VNFRDIPLYLQRKIQKAPMGSLSEKAAQKLLEDEKHNKRQIDQSITDILR 389
Db 344 GRLETKMEVNDAELFLMQMQRSNHOSNEKTDILKQIAETKHKRHDGSVELIGV 403
Qy 390 LSQKOTVNLNLSTRITGQPLVDDCFKPLVNSFKHCGATVHYGLKYTGATLNICNM 449
Db 404 LLYPGKGSSVLSVRAPGSSLVDDDTCLSMVRFETHCGLTQGMKHMRAFNCS 463
Qy 450 GVDVKOTVSAIQAC 464
Db 464 GV----SEASMEAC 474

RESULT 12
T10944 cysteine proteinase (EC 3.4.22.-) precursor - spring vetch
C;Species: Vicia sativa (spring vetch, tare)
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 20-Jun-2000
A;Accession: A60145; PMID:90342941
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-429 <ELA>
A;Accession: B60145
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-429 <ELA>
A;Accession: B60145
A;Molecule type: protein

A; Residues: 32-57 <EL2>
 R; Davis, A.H.; Nanduri, J.; Watson, D.C.
 J. Biol. Chem. 262, 12851-12855, 1987
 A; Title: Cloning and gene expression of *Schistosoma mansoni* protease.
 A; Reference number: A27378; PMID:87308326; PMID:3305515
 A; Accession: A27378
 A; Molecule type: mRNA
 A; Residues: 77-309; 'W', 311-429 <DAV>
 A; Cross-references: GB:MT7423; NID:9161060; PID:9161061
 C; Superfamily: legumain
 C; Keywords: glycoprotein; hydrolase; proteinase
 F; 1-19/Domain: signal sequence #status predicted <SIG>
 F; 20-31/Domain: propeptide #status predicted <PRO>
 F; 32-291/Product: hemoglobinase #status predicted <MAT>
 F; 292-429/Domain: carboxyl-terminal propeptide #status predicted <CTP>
 F; 192,205-215/Binding site: carbohydrate (Asn) (covalent) #status predicted <MAT>

Query Match Score 790; DB 2; Length 429;
 Best Local Similarity 40.5%; Pred. No. 5.6e-47;
 Matches 186; Conservative 75; Mismatches 156; Indels 42; Gaps 13;

Qy 10 ILVFHALLIFSAERSRKTOILNDYEDSSKS-
 Db 8 LSTITHLRL-----KCQL-
 Qy 69 AYQILRKGGKLKDENNIVMDDIAFSSENPRGVYLINKPDGEDVYKGVPKDYTKAEAVNVO 128
 Db 60 AYHVLRSGKIGKPEHTITMYDDIAYNLMPFPGLFLNDYNHKDNYEGVVDYREKVNNEK 119
 Qy 129 NYFNVNLLGNESVTGNGKVVKSGPMNFIYADHGAPLAMPTGDEVMAKDFEVEVL 188
 Db 120 TFLKVLKGDKS---AGGVKLKGKNDDVPIFYDHGAPLAAFP-DDELYAKEFMSLTK 174
 Qy 189 KMHKIKYKNRMVIVYEAEGSMFILQLKRNIVIAYTAANSKSSWGYCPESYPPPPS 248
 Db 175 YLHSHKRYSKLVITYTEANEGSMMQQILSNLSIIVATTAANSTECSYSPFGD----P 228
 Qy 249 EIGTCLGDTFTSISNLEDSDLHDSKETLEQQHIVVKRGSDPETSHWCRTGTEKMLKD 308
 Db 229 TRTCIADLSYNNATVDQSQTHLTORLQQKEYKRET---DLSHVQRQYGTMRMGRLL 283
 Qy 309 YLSSYIGRNPENDNTFTESFSPIISNSGLVNPDIPLYLQRKIQKAMPGSLESKEAQK 368
 Db 284 YSEFQGSRDKS---SENDEPPMKPRAHSIASRDPILTHRQIMMT--NNADPKSFLM 337
 Qy 369 KLDDERINHRKQIDQSTTDLRSYKQTQNVNLNSTRGFLYDWDCEFKTLVNSFKHH 428
 Db 338 QIGLKLKRDL---IEDPMKLIVKVMNNEIPNPKATIDQTL---DCTESVYEQFKSK 390
 Qy 429 CGATVHYGLKYTG---ALANICMGVGVKOTVSAEQAC 464
 Db 391 C-FTLQQAPEVGHFSTLNYCADSYTAETINEAIKIC 428

RESULT 14
 T19231 probable cysteine proteinase (EC 3.4.22.-) T28H10_3, precursor [similarity] - Caenorhabditis elegans
 C; Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jun-2000
 C; Accession: T19231; T25439
 R; Dobson, R.
 A; Reference number: Z19094
 A; Accession: T19231
 A; Status: preliminary; translated from GB/EMBL/DBJ
 A; Molecule type: DNA
 A; Residues: 1-462 <WIL>
 A; Cross-references: EMBL:277653; PIDN:CA01126..1; GSPDB:GN00023; CESP:T28H10_3
 A; Experimental source: clone C13C12
 R; Kershaw, J.
 A; Reference number: Z20034
 A; Accession: T25439

RESULT 15
 S31908 hemoglobinase - fluke (*Schistosoma japonicum*)
 C; Species: Schistosoma japonicum
 C; Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jun-2000
 C; Accession: S31908
 R; Marckelbach, A.; Hasse, S.; Dell, R.; Eschbeck, A.; Ruppel, A.
 A; Description: cDNA sequences of schistosoma japonicum coding for hemoglobinase and t
 A; Reference number: S31907
 A; Accession: S31908
 A; Status: preliminary
 A; Molecule type: mRNA
 A; Residues: 1-423 <MER>
 A; Cross-references: EMBL:X70967; NID:911164; PID:911165
 C; Superfamily: legumain

Query Match Score 768; DB 2; Length 423;
 Best Local Similarity 38.4%; Pred. No. 1.8e-45;
 Matches 178; Conservative 70; Mismatches 167; Indels 48; Gaps 13;

Qy 8 FQILVFLHALLIFSAESRKTQLLNDDVESSDIKAGTRAWAVLAVGSNEYNNRHQADIC 67
 Db 2 FYSIFFHLRLRVLD - - - - - CNEYSEENVDRHK - - - - - WAVLAVGSNGFENYRHQADVC 52
 Qy 68 HAYQILRKGLKLDENIIVEMYDIAFSSENPREGVINPKDVTGVPKDVTKEAVNV 1.27
 Qy 53 HAYHVLLSKGVKPHITMYDIAHNKENPFGKTDNDYRHKDYYKGVVTDYKGKVNP 11.2
 Db 128 QNFVNVLGNESGTGGNGKUVVSGPNNTIYADGAPGLIAMPGDEWMAKDNEVL 187
 Db 113 KTEFLQVLKGDKR - - AGKVLSKGNDVFTYFTDAGPATI AFP DDDIHAKPINTL 167
 Qy 188 EKMHKRKKYNMVITYVEACESGSMFEGTLKKNUNIYAVTAANSKESSWGVYCPESYPPPP 247
 Qy 168 KYLROHRRTSKLVIVYACESEGSNFAGLPTDINIYATTAARDPESSYATPCDD - - - - - 221
 Db 248 SEIGTCLGDTFSISWLEDSDLHMSKETLEQQYHVVRRVGDPVETSHCRFGTEKMLK 307
 Qy 222 PRISSCCLADLYSTDWSEKHOQTORTLDQQYKEYKFEI - - - - - NLSHQRYGDKMGK 276
 Qy 308 DYLSYIGRNPNENDFTTESFSSPISNSGLYMPRDIPPLYQRKTOKAPGSLESKEAQ 367
 Db 277 LYLSSEFOGRKK - - ASTEHDPMPMKRDKSFSRDPHTLHRRTMMA - NMMDNRTLL 330
 Qy 368 KKLIDEKNRKQI - - DQSITDLRLSVQTNVNLLTSTRUTGQPLVDDNDCFKTLVNS 424
 Db 331 MKIUGLKLKRDLIKDTMEVIDEMFNVKQPN - - - - - SNATDET M - - - DCIEVYKE 380
 Qy 425 FKNICGATHYGYKRTGAA - - NICANGDVYKQTVSAIEQAC 464
 Db 381 FQSKC-FKIQQAPEITGYLSTLNYCQKGYSKENNINGVIRKVC 422

Search completed: May 27, 2003, 15:28:33
 Job time : 52 secs

RESULT 1						
VPE_VICSA	ID	VPE_VICSA	STANDARD:	PRT:	4 9 3 AA.	
AC P49044;						
DT 01-SEB-1996 (Rel. 33, Created)						
DT 01-SEB-1996 (Rel. 33, Last sequence update)						
DT 15-JUN-2002 (Rel. 41, Last annotation update)						
DE Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE) (Proteinase B).						
OS <i>Vicia sativa</i> (Spring vetch) (Rare)						
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Speciamotophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Vicieae; <i>Vicia</i> .						
OC NCBI_TAXID=3908;						
OX [1]						
RN RP SEQUENCE FROM N.A.						
RC TISSUE=Seed;						
RX MEDLINE=95220376; PubMed=705362;						
RA Becker A.D., Shutov A.D., Nong V.H., Senyuk V.I., Jung R., Horstmann C., Fischer J., Nielsen N.C., Muntz K.;						
RA "Purification, cDNA cloning, and characterization of protease B from germinating 'vetch seeds'";						
RT Biochimia 47:814-811(1982).						
RT asparagine-specific endopeptidase from germinating vetch (<i>Vicia sativa</i> L.) seeds.";						
RL Eur. J. Biochem. 228:456-462(1995).						
RN [2]						
RP CHARACTERIZATION.						
RC TISSUE=Seed;						
RX MEDLINE=92232729; PubMed=7046813;						
RA Shutov A.D., Do N.L., Vaintraub I.A.;						
RA "Purification and partial characterization of protease B from germinating 'vetch seeds'";						
RT Biochimia 47:814-811(1982).						
CC -!- FUNCTION: ASPARAGINE-SPECIFIC ENDOPEPTIDASE INVOLVED IN THE PROCESSING OF VACUOLAR SEED PROTEIN PRECURSORS INTO THE MATURE FORMS.						
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.						
CC This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation						
CC the European Bioinformatics Institute. There are no restrictions on its						
CC use by non-profit institutions as long as its content is in no way						
CC modified and this statement is not removed. Usage by and for commercial						
CC entities requires a license agreement (see http://www.isb-sib.ch/announce) or send an email to license@isb-sib.ch).						
CC -						
CC DR EMBL; 734899; CAA84383..1; -						
DR MEROPS; C13.002; -						
DR InterPro; IPR001096; Legumain.						
DR Pfam; PF01650; Peptidase_C13; 1.						
DR PRINTS; PR0077; HEMOGLOINASE.						
KW Hydrolase; Thiol protease; Signal.						
FT SIGNAL 1 193 POTENTIAL.						
FT CHAIN 20 193 VACUOLAR PROCESSING ENZYME.						
FT ACT_SITE 174 174 POTENTIAL.						
FT ACT-LITE 216 216 S -> L.						
FT VARIANT 88 88						

FT VARIANT	89	89	K -> R.	CC MAXIMUM LEVELS IN FULLY COLORED FRUIT. THE LEVELS INCREASE DURING FLOWER DEVELOPMENT AND SHOW HIGHEST LEVELS IN FLOWERS AT ANESTHESIA.
FT VARIANT	228	228	D -> E.	CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.
FT VARIANT	254	254	P -> S.	CC -----
FT VARIANT	366	366	R -> H.	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. CC use by non-profit institutions is granted on the understanding that no CC modification and thus statement is removed. Usage by and for commercial CC entities requires a license agreement (see http://www.isb-sib.ch/announce/ CC or send an email to license@isb-sib.ch).
SQ SEQUENCE	493 AA:	54383 MW:	FEIFF8ADAFBB63AE CRC64;	CC -----
Query Match	51.7%	Score 1268.5; DB 1; Length 493;	CC -----	CC -----
Best Local Similarity	53.7%	Pred. No. 8.2e-83; Gaps 3;	CC -----	CC -----
Matches	235;	Mismatches 70; Indels 11; Gaps 3;	CC -----	CC -----
Qy 31	NDNDYBESSDKSAKGRTRWAVLVAAGSNSEYYNRHOAIDCHAYQIYLRRGGLDENIVMFYDD 90	FT SIGNAL	1 20	CC -----
Db 48	NNDFDF-----GTRWAILLAGSGNGWYWRHQSDVCHAYQIYLRRGGSKEEENIVMFYDD 100	FT CHAIN	21 494	CC -----
Qy 91	IAFFSENPRPQVIINKPKDGEDPVYKVPKDVTKEAVNQMEYAVLGNNSGVTGGNGKVKK 150	FT CARBOHYD	151 151	CC -----
Db 101	IASNEDNPRAPIVINEKPGDGDVYACPVKTYGAVHADNFYAALGLNSALTGSGRKVD 160	FT CARBOHYD	336 336	CC -----
Qy 151	SGPDNDNFIYIYADHGAPGLIAMPTEGDEMAKDFNEVLEKMMHKRKYKYNMIVYFACESGS 210	FT ACT SITE	178 178	CC -----
Db 161	SGPDNDHIFVYIYTGHGVLGVMPGPYASLDLNEVLKRKHASTCYKSLSVFLYFACESGS 220	FT ACT SITE	220 220	CC -----
Qy 211	MFGELPLKKRNNAATTAANKESSMVGCVPESESPPPSERGTGCTGDTTSISWEDSLHD 270	FT SEQUENCE	494 AA: 54291 MW: 8B2059237E03B2B8.CRC64;	CC -----
Db 221	IFEGGLELPDDLNIVYATASNAESESSMGYICPGDKKPPPPYSTICGLDLYSIAMWEDSEFHN 280	Query Match	51.4%; Score 1262.5; DB 1; Length 494;	CC -----
Qy 271	MSKETELBQOYHVVKRVRGSDPTEHVCRGTERMLKDLYSSLSSYGRNENDNETTESFS 330	Best Local Similarity	54.6%; Pred. No. 2.2e-82;	CC -----
Db 281	LQTESUQSKQKLVKURKRTIE-PYCSHVMYGDIQSKLNDYQIIGTNANDNSFVDETE 339	Matches	69; Mismatches 120; Indels 9; Gaps 3;	CC -----
Qy 331	SPI---SNSGLVNPDRDPLYLQRKIQAPMGSLSEKAQKLLDEKMRKQDQSITDI 387	Qy 30	UNDNDYEVSSDSRSAGKTRWAVLVAGSNSEYYNRHOAIDCHAYQIYLRRGGLDENIVMFYDD 89	CC -----
Db 340	NSLKLTPTSAAVNQRDADLTHFWENFKRPAEGPSSQKNEAEKQVLEAMSMRKHIDNSYRLI 399	Db 50	VNDDD-----DSVGTTRWAVLLAGSNGFWNNYRHOAIDCHAYQIYLRRGGLDENIVMFYDD 103	CC -----
Qy 388	LRLSVKQTNVNLNLSTRTRGQPLYDDWDCFCFTVNFSKHNHGATVHYGLKYTGALANIC 447	Qy 90	DAFSENPRQVIAKTPDGEDVYKGPKDVTKEAVNQMEYAVLGNNSGVTGGNGKV 149	CC -----
Db 400	GQLFGTIEKGTELLDVVRPAGSPSLVDNWDCLTKNTKTFETHCGSLSQYGMKHMRSFANIC 459	Db 104	DAFNEENPRQVITNHPHGDVYKGPKDVTGEDETVKFAVULGNKTAALTGSKRVV 163	CC -----
Qy 448	NMGVDYKQTQVSAAIOACS 465	Qy 150	KSGPNDNFIYIYADHGAPGLIAMPFGEDEWAKDFNEVLEKMHKRRKVKNMVIVYEAESGS 209	CC -----
Db 460	NAGIPNEPMEEAQACA 477	Db 164	DSPGNDHIFIYSDHGPVGGMPTSRYIATEDELIDVLAKKHASGNYKSLVFLYEAESGS 223	CC -----
Qy	SMFEGTIKKNNIYATAANKESSMVGCVPESESPPPSERGTGCTGDTTSISWEDSLHD 269	Qy 210	SMFEGTIKKNNIYATAANKESSMVGCVPESESPPPSERGTGCTGDTTSISWEDSLHD 269	CC -----
Db	SIEFGLLLEGNLNITYATTASNAESESSMGYICPGDKKPPPPYSTICGLDLYSIAMWEDSEFH	Db 224	SIEFGLLLEGNLNITYATTASNAESESSMGYICPGDKKPPPPYSTICGLDLYSIAMWEDSEFH 283	CC -----
Qy	DMSKETELBQOYHVVKRVRGSDPTEHVCRGTERMLKDLYSSLSSYGRNENDNETTESFS 329	Qy 270	DMSKETELBQOYHVVKRVRGSDPTEHVCRGTERMLKDLYSSLSSYGRNENDNETTESFS 329	CC -----
Db	NAETETHQOYELVTRTAYNSCYSHVNQYDGIGLSKSNLFTYLGTPANDNTYTFVDE	Db 284	NAETETHQOYELVTRTAYNSCYSHVNQYDGIGLSKSNLFTYLGTPANDNTYTFVDE 343	CC -----
Qy	S-SPINSGLVNPDRDPLYLQRKIQAPMGSLSEKAQKLLDEKMRKQDQSITDI 388	Qy 330	S-SPINSGLVNPDRDPLYLQRKIQAPMGSLSEKAQKLLDEKMRKQDQSITDI 388	CC -----
Db	SURPASKA--VNORDADLHEWDYKRPAGEGTPRKAEOKFAMSHMHVHSIKLIG 401	Db 344	SURPASKA--VNORDADLHEWDYKRPAGEGTPRKAEOKFAMSHMHVHSIKLIG 401	CC -----
Qy	RLSVKGQTNVNLNLSTRTRGQPLYDDWDCFCFTVNFSKHNHGATVHYGLKYTGALANIC 448	Qy 389	RLSVKGQTNVNLNLSTRTRGQPLYDDWDCFCFTVNFSKHNHGATVHYGLKYTGALANIC 448	CC -----
Db	KLFGEKGPPLNTVRPAGPLVDDWGCLKSLVRFESHGALSQYGMKHMRSLANIC 461	Qy 449	MGVDYKQTQVSAAIOACS 464	CC -----
Qy	TCIGKEKMAEASQAQC 477	Db 462	TCIGKEKMAEASQAQC 477	CC -----
RESULT 2				
VPE_CITSI	STANDARD;	PRT:	494 AA.	RESULT 3
ID	VPE_CITSI			ID VPEG_ARATH STANDARD;
AC	P49043;			AC Q39119; O95B73;
DT	01-FEB-1996 (Rel. 33, Last sequence update)			DT 16-OCT-2001 (Ref.: 40, Created)
DT	01-FEB-1996 (Rel. 33, Last annotation update)			DT 16-OCT-2001 (Ref.: 40, Last sequence update)
DE	Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE).			DE Vacuolar processing enzyme, gamma-isozyme precursor (EC 3.4.22.-)
OS	Citrus sinensis (Sweet orange).			DE (Gamma-VPE)
OC	Spermatophyta; Streptophyta; Embryophyta; Tracheophyta;			DE (Gamma-VPE)
OC	Magnoliophyta; eudicots; core eudicots; Rosidae;			GN AT4G32940 OR F26P21.60.
NCBI TaxID	2711;			
RN	SEQUENCE FROM N.A.			
RP	Strain=cv. Washington Navel; TISSUE=Flavedo;			
RX	MEDLINE=96030254; PubMed=7483446;			
RA	Alonso J.M., Granell A.;			
RT	"A putative vacuolar processing protease is regulated by ethylene and also during fruit ripening in Citrus fruit."			
RL	Plant Physiol. 109:541-547 (1995).			
CC	FUNCTION: ASPARAGINE-SPECIFIC ENDOPEPTIDASE THAT MAY BE INVOLVED IN PROCESSING OF PROTEINS TARGETED TO VACUOLE AS FLOWER OPENING AND FLAVEDO DECREASING.			
CC	-!- TISSUE SPECIFICITY: HIGH LEVELS ARE SEEN IN THE FLOWERS, A LOWER LEVEL EXPRESSION IS SEEN IN THE LEAVES, WHILE VERY LOW LEVELS ARE SEEN IN THE STEMS AND ROOTS.			
CC	-!- DEVELOPMENTAL STAGE: THE LEVELS ARE LOW IN GREEN FRUITS, BUT ACCUMULATE WITH COLOR CHANGE OCCURRING DURING RIPENING, REACHING			

Arabidopsis thaliana (Mouse-ear cress).
 Eukarya; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicots; core eudicots; Rosidae;
 eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI_TaxID=3702;

[1]
 SEQUENCE FROM N.A.
 STRAIN=cv Columbia;
 MEDLINE=96146075; pubmed=8589932;
 Kinoshita T., Nishimura M., Hara Nishimura I.;
 "The sequence and expression of the gamma-YPE gene, one member of a
 family of three genes for vacuolar processing enzymes in *Arabidopsis*
thaliana";
 Plant Cell Physiol. 36:1555-1562(1995).

[2] SEQUENCE FROM N.A.
 STRAIN=CV Columbia;
 MEDLINE=20083488; PubMed=10617198;

Pohl T., Schueler C., Wambout R., Murphy G., Volckaert G.,
 Entian K.-D., Terry N.,
 Harris B., Ansorge W., Brandt P., Grivell L.A., Weichselgartner M., de Simone V., Obermaier B., Mache R.M., Mueller M.,
 Kreis M., Delseny M., Puigdomenech P., Watson M., Schmidtneini T.,
 Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,
 Vos P., Hoheisel J., Zimmermann W., Wiedler H., Ridley P.,
 Langham S.-A., McCullagh B., Blaham L., Robben J.,
 van der Schueren J., Gromyronpr B., Chuang Y.-J., Vandenbussche F.,
 Braeck M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoort E.,
 Weitzenegger T., Bothé G., Ramsperger U., Hilbert H., Braun M.,
 Holzer E., Brandt A., Peters S., van Staveren M., Dirkse W.,
 Moijman P., Klein Lankhorst R., Rose M., Hauf J., Koetter P.,
 Berneiser S., Hempel S., Feldpausch M., Lambertz S., Van den Daele H.,
 De Keyser A., Buysshaert C., Gielen J., Villarreal R., De Clercq R.,
 Van Montagu M., Rogers J., Cronin A., Quail M., Bray Allin S.,
 Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,
 Pettett A., Rajandream M.A., Lyne M., Benes V., Reichmann S.,
 Borkova D., Bloecker H., Schärfe M., Grimm M., Lehnert T.-H.,
 Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,
 Gabel C., Fuchs M., Fartmann B., Grandjean K., Dauner D., Herzl A.,
 Neumann S., Argilio A., Vitalte D., Liguroi R., Piravandi E.,
 Massenet O., Quigley F., Clabaud G., Muendlein A., Felber R.,
 Schnabl H., Hiller R., Schmidt W., Lecharny A., Aubourg S.,
 Chefdor F., Cooke R., Berger C., Monfort A., Casacuberta E.,
 Gibbons T., Weber N., Vandembol M., Bargues M., Terol J., Torres A.,
 Perez-Perez A., Purneille B., Bent E., Johnson S., Tacon D., Jesse T.,
 Heinen L., Schwarz S., Scholler P., Heber S., Francis P., Bielke C.,
 Prishman D., Haase D., Lemcke K., Mewes H.W., Stocker S.,
 Zaccaria P., Bevan M., Wilson R.K., Schulz K., Habermann K.,
 Parnell L., Dedhia N., Gnoj L., Huang E., Spiegel L.,
 Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threideh J.,
 Stoneking T., Kalicki J., Graves T., Harmon G., Edwards J.,
 Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,
 Minx P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,
 Kramer J., Fulton L., Mardis E., Dante M., Pepin K., Hillier L.,
 Nelson J., Spilker J., Ryan F., Andrews S., Geissel C., Lawrence D.,

Du H., Ali J., Berghoff A., Jones K., Drone M., Joshi C., Antoniou B., Zidiani M., Strong C., Sun H., Lamar B., Yordan C., Ma P., Zhong J., Preston R., Vil D., Shekher M., Matero A., Shah R., Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Till S., Granat S., Shohdy N., Hasegawa A., Haneed A., Lodi M., Johnson A., Chen E., Marras M., Martienssen R., McCormie W.R.; "Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana"; Nature 402:769-777(1999).
 -!- FUNCTION: ASPARAGINE-SPECIFIC ENDOPEPTIDASE INVOLVED IN THE PROCESSING OF VACUOLAR SEED PROTEIN PRECURSORS INTO THE MATURE FORMS (BY SIMILARITY).
 -!- SUBCELLULAR LOCATION: vacuolar.
 -!- TISSUE SPECIFICITY: SPECIFIC TO VEGETATIVE ORGANS.
 -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

Query	Match	50 %	Score 1246	DB 1;	Length 490;
Best Local Similarity	53.6%	Pred. No. 3.2e-81;			
Matches	233;	Conservative	71;	Mismatches	121;
Gaps	10;				
Dy	36 ESSDKSAKGRWAVLYAGSNSEYYNRHQADICHAYQILRGKGGLKDENITVEMDDIAFS 95				
Db	45 ENDDDSNSTRWAIVAGSGYWNTRHQADICHAYQILRGKGLEENITVEMDDIANNY 104				
Dy	96 ENPRPGVIINKPGDGYVKGPDKYTKEAVNVQNFYINVLLNGESVTGNGKVVKSGPND 155				
Db	105 ENPRPGTINSPHGRDGVYQGPDKYTGFIDVNUFLNLFATVQGSGKVVKSGPND 164				
Dy	156 NIFTYADHGAPKGTLAMPTDEWMAKDFNEVILEXNMHRKRKYNAKTVIYACESSGMPE 215				
Db	165 HIFIFYSDHGGPVLGMPSPYLANDLVDLKHKHALGTYSKSLVYLEACESGSSTIFGL 224				
Dy	216 LKKNNIYAVTAANSKESSMGVYCEPPSEIGTCGDTFTSISWLEDSDLHDMSKET 275				
Db	225 LPEGNIYATAASNEAESSNGTYCPEEEPSPPPETCTGDLTSVAMMDSGMHNLOPTE 284				
Dy	276 LEQQYHVKR--NGSDVPETSHVCRCGETKMLDKDLYSSIGNPNDNFTPE--SFS 330				
Db	285 LHQQBLVKRRTAPGYSY -GSHMVOYGDVGISKRDNLQDLYMTGNPANDEFAADNSL 342				
Dy	331 SPISNSGLVNFRDIPYLILORKIKAPMGSLESKAQOKKLDEENHRKQIDQSFTDILRL 390				
Db	343 PP--SRVTNQRDADLVHWEKYKAPESSARKTEAQKYLEAMSHRLJHDNSVILVSKJ 399				
Dy	391 SVKQTQVNLLJSTRTGQPLDVMDCFKLTVNSFKNHGATVHYGLKTYGALANICNG 450				
Db	400 LFGISRGPEVLNKVRNSQGQFLVDWNCNLQVRAFERHCGLSOYGIKMRSPFANICNG 459				
Dy	451 VDVKOTVSALEQAC5 465				

RESULT 4
 VPEA_ARATH STANDARD; PRT; 478 AA.
 ID VPEA_ARATH
 AC P49047; 082806;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vacuolar processing enzyme, alpha-isozyme precursor (EC 3.4.22.-)
 DE (Alpha-VEP).
 GN AT2G2940 OR T19L18.25.
 DS Arabidopsis thaliana (Mouse-ear Cress).
 DS Bokarvaya; Viridiplantea; Streptophyta; Embryophyta; Tracheophyta;
 DS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 DS eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 NCBI TaxID=3702;

RL Submitted (OCT-2001) to the EMBL/GenBank/DDJB databases.

CC -1- FUNCTION: ASPARAGINE-SPECIFIC ENDOPEPTIDASE INVOLVED IN THE
CC PROCESSING OF VACUOLAR SEED PROTEIN PRECURSORS INTO THE MATURE
CC SUBCELLULAR LOCATION: Vacuolar.

CC -1- TISSUE SPECIFICITY: SPECIFIC TO SEED.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.

CC -1- CAUTION: Ref. 2 sequence differs from that shown due to erroneous
CC gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC DR D61394; BAA9615_1; -;

DR EMBL; ACG007190; AF019550_1; ALT_SEQ.

DR EMBL; AV059156; APL15381_1; -;

DR MEROPS; C13_001; -;

DR InterPro; IPR00106; Legumain.

DR Pfam; PF01650; Peptidase_C13_1.

DR PRNTS; PR00776; HEMOGLOBINASE.

KW Hydrolase; Thiol Protease; Signal; Multigene family.

FT SIGNAL 1 21

FT ACT_SITE 22 486

FT VACUOLAR PROCESSING ENZYME, BETA-ISOZYME.

FT CHAIN 221

FT MISSING (IN REF. 1).

FT CONFLICT 288

FT ACT_SITE 211

FT D -> E (IN REF. 1).

FT CONFLICT 324

FT V -> A (IN REF. 1).

FT CONFLICT 337

FT Y -> H (IN REF. 1).

FT CONFLICT 358

FT 486 AA; 53828 MW;

SQ SEQUENCE 8B3EFD3CDABC9C3 CRC64;

Query Match 47.3%; Score: 1160; DB: 1; Length: 486;
Best Local Similarity 49.3%; Pred. No. 4.le-75;
Matches 232; Conservative 68; Mismatches 149; Indels 22; Gaps 7;

Qy 7 HQQILVFLHALLFSAESR-----KTOQLINDNVESSDIAKTRAVLVAGSNYE 57

Db 6 YFRPALLLVLVLYHAESRGREFPKLMPTEANPAP---QDEDVGTRAVLVAGSSGY 62

Qy 58 YNRHQADITCHAQYLRLKGGLKDENIIVMFYDDIAFSENPREGVVINKPDGEDVYKGPV 117

Db 63 GNVRHQADIVCHAWQYLRLKGGLKEENIVLMDYDIANHPLNPRTGLNHPDDDTAVGP 122

Qy 118 KDTIKEAIVNQVNYVNLGNESVTGGNGKVKVSGPNDNIFYYADHGAPGLAMPGDE 177

Db 123 KDTGSSVTAANQYAVLGDQKAVKGGSKVIAKSPNDHFITYADHGPGVLMGPNTPH 182

Qy 178 VMKDFNEVLEKHKRKYKVNKMYVYEAESGSMFEGTILKUNNIYVATAANSKESSWGV 237

Db 183 IYADFIETLKKHASSTYKEVWIVYEAESGSIFEGIMPKDLNIVYTASNAQESSYGT 242

Qy 238 YCPESYPPPSEIGTCLGDTFSTSWLEDSDLHDMSKETLEQOYHVVYKRRVG -SDVPETS 295

Db 243 YCGMNPSPPEWITCGDLYSAWMDSETHLKETIKQOHTVMTSNNNTYSGG 302

Qy 296 HYCRFGTEKMLKDYLSSYIGRNPPENDNFTFESFSSPI -SNSGLVNPRDIPLYLQRKIQ 354

Db 303 HYMEYGNNSIKSEKLYLQGDFATVNLPLNE --LEVKSKitGVNNORDADLFLWHMYR 359

Qy 355 KAFMGLESKEAQKLDEKHNHQKQDOSITDILRSVQQT -NVNLJLTSTPTGQPLVD 413

Db 360 TSEDGSRKKDDT1KELTETTRIRKHLDASVELIATLFGPTVNVLV-- -REPGLPLVD 416

Qy 414 DNQCFKTLYNSPNFHCGATVHYGLKTYGALANCNMGVDVKOTVSATEQAC 464

Db 417 DNWECLKSARVRFEEHCGSLTQGMKMRFAFNCCNNGVSKELMEAATAC 467

RESULT 6

VPE_RICCO STANDARD; PRT; 497 AA.

ID VPE_RICCO

AC P4942;

DT 01-FEB-1996 (Rel. 33; Created)

DT 01-FEB-1996 (Rel. 33; Last annotation update)

DT 16-OCT-2001 (Rel. 40; Last annotation update)

DE Vacuolar processing enzyme precursor (EC 3.4.22.-) (VPE).

OS Ricinus communis (Castor bean).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC euroids I; Malpighiales; Euphorbiaceae; Ricinaceae; Ricinus.

OX NCBI_TaxID=3988;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-94146557; PubMed=8312744;

RA Harai-Nishimura I.; Takeuchi Y.; Nishimura M.;

RT "Molecular characterization of a vacuolar processing enzyme related to a putative cysteine proteinase of *Schistosoma mansoni*.";

RT Plant Cell 5:1651-1659(1993).

RN [2]

RP CHARACTERIZATION.

RX MEDLINE-92077151; PubMed=1743299;

RA Harai-Nishimura I.; Inoue K.; Nishimura M.;

RT "A unique vacuolar processing enzyme responsible for conversion of several proprotein precursors into the mature forms.";

RT FEBS Lett. 294:89-93(1991).

CC -1- FUNCTION: ASPARAGINE-SPECIFIC ENDOPEPTIDASE INVOLVED IN THE
CC PROCESSING OF VACUOLAR SEED PROTEIN PRECURSORS INTO THE MATURE
CC FORMS

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC DR MEROPS; C13_001; -;

CC DR InterPro; IPR01086; Legumain.

DR PRINTS; PR00776; HEMOGLOBINASE.

DR Hydrolase; Thiol protease; Signal.

FT SIGNAL 1 31

FT CHAIN 32 497

FT ACT_SITE 180 180

FT ACT_SITE 222 222

FT POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 180 180

CC DR BAA04225_1; -;

CC DR DR PRINTS; PR00776; Peptidase_C13; 1.

CC DR FT ACTSITE; Thiol protease; Signal.

CC FT SIGNAL 1 31

CC CC VACUOLAR PROCESSING ENZYME.

CC CC ACTSITE 222 222

CC CC POTENTIAL 1

Qy	275 TLEQQYHVKRKEVG--SDVPETSHYCRFGTEKMLKDYLSSYIGRNPENDNETTESFSSP	332	FT PROTEIN	16	35	LEGUMAIN.
Db	291 TYKQYQSSVKARTSNVNTYANGSHM/QYGNOSTIKRDKLYLFOGDPASNF----P	342	FT CHAIN	36	475	N-LINKED' (GLCNAC. . .) (POTENTIAL).
Db	333 ISNGL----VNPRDIPLYLQRKIQKAPMSLESKQKKLDEKNHRKOIDQSITID	386	FT CARBOHYD	300	300	POTENTIAL.
Qy	343 PNAHLUNAPMEVYNQDAELHFMWOLYKRSENGSEKKETLQOQTDAIKHRSHIDSSHSQL	402	FT ACT-SITE	158	158	POTENTIAL.
Db	387 ILLRLSVKOTNVNLNLSTRTGQPLVDDWDCFKTLYNSFKNCATVHYGLKTGALANI	446	FT ACT-SITE	200	200	POTENTIAL.
Qy	403 IGDLLFEPKPKASSAUKSRSVREPGSPVLDNGCLSKMVMYRFETCCSSLTOYGMKHMTFANI	462	SEQUENCE	475 AA:	52763 MW:	6D1C1D6B72C5504C CRC64;
Qy*	447 CNMGVDYKQTVSAIEQACS	465	Query Match	44 / 0%	Score 1080;	DB 1;
Db	463 CRAGVSHTSMSMEACNACCS	481	Best Local Similarity	48 / 3%	Pred. No. 1.	Length 475;
Db	98 INHPOGPDVAYAVGPKDYGTVENVTPELYAVLGDKSKVKGSGRVINSNPEDRIFTYSD	157	Matches 207; Conservative	73;	Mismatches 129;	Indels 20;
Qy	104 TNPQDGDDVYGVPKDYTEKAVVNQFYNVNLGNSGVYGTGGVVKVSKPNDFNITYAD	163	Db 38 GTRWAVLVAGSGNGYRNHQDVCAYQOLLKGKVKEENIVFMYDDAYNAAMNPRPGV	97		
Qy	111 INHPOGPDVAYAVGPKDYGTVENVTPELYAVLGDKSKVKGSGRVINSNPEDRIFTYSD	157				
Qy	164 HCAGPLIAMPQDGEVMAKDFNEVLEKMHMKRKYKVNKMVIYEAESGSMPFGLKKLNLY	223				
Db	158 HCGPGVIGMPNAPPEYAMDFDVLEKKKHAGGYKDNVITYTEACSGSMPFGLKKLNLY	217				
Qy	224 AVTAANKESSWGVYCPESYPFPPEIGTCGDTFSISNLEDSDLHDMSKTELQEQYHVV	283				
Db	218 VTTAQNQENSTGTCGPMNPPPEEYTCGLDLYVSVMEDSETHNLKRETVOOTOSV	277				
Qy	284 KERVGSDVPEH-SHVCRFGTEKMLDYLSYYGRNPENDFTFESFSPISNSGL--	338				
Db	278 RAKTSNSNSYRGSHMNOYGTNTIAEKLYLHGDPATVNF-----PHINGNLREAK	329				
Qy	339 --VNPRDIPLLYQKTKQAPMGSLSESQAQKLDEKHKQIDQSTIDLRLSVKOT	395				
Db	330 MEVVNQNDAALELMKOMYQRNHQPKKTHLQEQTETYKRNHLDGSELVGLYLGPG	389				
Qy	396 NVLNJLTSTRITGQPVODNCFKTLVNSENKHNGCATVHYGLKTYGALANICNMGYDKQ	455				
Db	390 KSSSVLHSVRAPGLPLVDDWTCLSKMVRFETHCSLRQGMKHMRAFNCNSGV---	445				
Qy	456 TVSAIBOAC	464				
Db	446 SKASMBEAC	454				
<hr/>						
RESULT 8						
VPE_SOYBN	ID	VPE_SOYBN	STANDARD;	PRT;	495 AA.	
AC	AC	P49045;				
DT	DT	01-FEB-1996 (Rel. 1.	33, Created)			
DT	DT	01-FEB-1996 (Rel. 33, Last sequence update)				
DT	DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	DE	Legumain precursor (EC 3.4.20.34) (Asparaginyl endopeptidase).				
Canavalia ensiformis (Jack bean) (Horse bean).						
Eukaryota; Viridiplantae; Streptophytina; Embryophytina; Tracheophytina; Spermatophytina; Magnoliophytina; eudicots; core eudicots; Rosidae; OX						
NCBI_TAXID=3823;						
RN	RN	SEQUENCE FROM N_A.				
RC	RC	TISSUE=Seed;				
RX	RX	MEDLINE=9515263; PubMed=7852272;				
RA	RA	Takeda C., Miura Y., Mitta M., Matsushita H., Kato I., Abe Y., Yokosawa H., Ishii S.;				
RT	RT	"Isolation and analysis of cDNA encoding a precursor of Canavalia ensiformis asparaginyl endopeptidase (legumain)." ;				
RT	RT	J. Biochem. 116:541-546 (1994).				
RN	RN	{2}				
RP	RP	SEQUENCE OF 36-60, AND CHARACTERIZATION.				
RX	RX	MEDLINE=9315205; PubMed=8429028;				
RA	RA	Abe Y., Shirane K., Yokosawa H., Mitta M.,				
RA	RA	"Asparaginyl endopeptidase of jack bean seeds. Purification, characterization, and high utility in protein sequence analysis." ;				
RL	RL	J. Biol. Chem. 268:3525-3529 (1993).				
RN	RN	13]				
RP	RP	REVIEW.				
RX	RX	MEDLINE=95147717; PubMed=7845236;				
RA	RA	Ishii S.; Legumain: asparaginyl endopeptidase."				
RT	RT	"Asparaginyl endopeptidase of jack bean seeds. Purification, characterization, and high utility in protein sequence analysis." ;				
RL	RL	Math. Enzymol. 244:604-615 (1994).				
CC	CC	-1-CATALYTIC ACTIVITY: Hydrolysis of proteins and small-molecule substrates at -Asn- l-Xaa-bonds.				
CC	CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.				
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	CC	INTERPRO: IPR01096; Peptidase C13; 1.				
DR	DR	EMOPS; C13-001; InterPro: IPR01096; Legumain.				
DR	DR	Pfam: PF01650; Peptidase C13; 1.				
DR	DR	PRINTS; PRO0776; HEMOGLOBINASE.				
KW	KW	Hydrolase; Thiol protease; Signal; Zymogen.				
FT	FT	POTENTIAL.				
<hr/>						
CC	CC	SEQUENCE FROM N_A.				
RC	RC	TISSUE=Seed Cotyledon;				
RX	RX	Medline=94156350; PubMed=8075902;				
RA	RA	Shimada T., Hirawa N., Nishimura M., Hara-Nishimura I.;				
RA	RA	"vacuolar processing enzyme forms"; RT the corresponding mature forms;"				
RT	RT	Plant Cell Physiol. 35:717-718 (1994).				
CC	CC	-1- FUNCTION: ASPARAGINE-SPECIFIC ENDOPEPTIDASE INVOLVED IN THE PROCESSING OF VACUOLAR SEED PROTEIN PRECURSORS INTO THE MATURE FORMS.				
CC	CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.				
CC	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its				

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC	endopeptidase";	
CC	modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	[2]	
CC	SEQUENCE FROM N.A.	
CC	RN [2]	
DR	RT RL RP	
DR	MEROPS; C13_001; - .	
DR	RA RX PubMed=9821970;	
DR	InterPro; IPR001096; Legumain.	
DR	RA "Autocatalytic activation of human legumain at aspartic acid residues";	
DR	RT residues";	
DR	RPLB Lett. 438:114-118(1998).	
DR	RN [3]	
DR	RP MUTAGENESIS OF HIS-47; CYS-152; HIS-150 AND CYS-191.	
FT	RX PMID=9891971;	
FT	RA Chen J.-M., Rawlings N.D., Stevens R.A., Barrett A.J.;	
FT	RT "Identification of the active site of legumain links it to caspases, cathepsin and granzipains in a new clan of cysteine endopeptidases.";	
FT	RT FEBS Lett. 441:361-365(1998).	
FT	RT -!- FUNCTION: Has a strict specificity for hydrolysis of asparagine bonds. Can also cleave aspartyl bonds slowly, especially under acidic conditions. May be involved in the processing of proteins for MHC class II antigen presentation in the lysosomal/endosomal system (By similarity).	
FT	RL CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins and small-molecule substrates at -Asn-1 Xaa-Bonds	
FT	CC -!- SUBCELLULAR LOCATION: Lysosomal.	
FT	CC -!- TISSUE SPECIFICITY: Ubiquitous. Particularly abundant in kidney and placenta.	
FT	CC -!- PRM: Glycosylated (Probable).	
FT	CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.	
CC	CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).	
CC	CC -!- DR EMBL: AJ000990; CAA04439; 1; -	
CC	DR EMBL: AF04266; AAF21659; 1; -	
CC	DR MEROP; C13_004; -	
CC	DR MGD; MGI:1330838; Legumain.	
CC	DR InterPro; IPR01096; Legumain.	
CC	DR PFAM; PF01650; Peptidase_C13_1.	
CC	DR PRINTS; PR00776; HEMOGLOBINASE.	
CC	KW Hydrolase; Thiol protease; Lysosome; Glycoprotein; Zymogen; Signal.	
CC	FT SIGNAL 1 17 BY SIMILARITY.	
CC	FT CHATN 18 325 LEGUMAIN.	
CC	FT PROPEP 326 435 BY SIMILARITY.	
CC	FT CARBOHYD 93 93 N-LINED (GLCNAC. . .) (POTENTIAL).	
CC	FT CARBOHYD 169 169 N-LINED (GLCNAC. . .) (POTENTIAL).	
CC	FT CARBOHYD 265 265 N-LINED (GLCNAC. . .) (POTENTIAL).	
CC	FT CARBOHYD 274 274 N-LINED (GLCNAC. . .) (POTENTIAL).	
CC	FT ACT-SITE 150 150 POTENTIAL.	
CC	FT ACT-SITE 191 191 POTENTIAL.	
CC	FT SITE 325 326 CLEAVAGE (AUTO-).	
CC	FT MUTAGEN 47 47 H->A: 54 % LOSS OF ACTIVITY.	
CC	FT MUTAGEN 52 52 C->S: NO LOSS OF ACTIVITY.	
CC	FT MUTAGEN 150 150 H->A: COMPLETE LOSS OF ACTIVITY.	
CC	FT MUTAGEN 191 191 C->S: COMPLETE LOSS OF ACTIVITY.	
SQ	SQ SEQUENCE 435 AA: 49373 MW; F955B9E10098013D CRC64;	
Query Match	Query Match 34.4%; Score 843; DB 1; Length 435;	
Best Local Similarity	Best Local Similarity 41.6%; pred; No. 1.2e-52;	
Matches	Matches 184; Conservative 70; Mismatches 138; Indels 50; Gaps 12;	
Qy	Qy 39 DKSAGKTRHAWAVGSENEYNYRHOAIDCHAYQILRKGGKLDENITIVMFDDIASSEN P 98	
Db	Db 24 DPEDGKHKWVWVQSGNGWVYQDQTLIPDEQITVMDDIANSEN P 83	
Qy	Qy 99 RPGVIVNRPNPTDVYKVLKDYGEDVTPPENFLAVLRGDPEAVKGKGSKVLSGP RDHV 157	
Db	Db 84 TPGVVINRPNPTDVYKVLKDYGEDVTPPENFLAVLRGDPEAVKGKGSKVLSGP RDHV 143	
OC	"Cloning and expression of mouse legumain, a lysosomal	
OC	Mus musculus (Mouse).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
NCBI_TaxID	NCBI_TaxID=10090;	
RN	RN [1]	
OC	SEQUENCE FROM N.A., AND CHARACTERIZATION.	
OC	MEDLINE=98A16120; Pubmed=9742219;	
OC	Chen J.-M., Dando P.M., Stevens R.A.E., Fortunato M., Barrett A.J.;	
RT	RT	

"Cloning, isolation, and characterization of mammalian legumain, an asparaginyl endopeptidase." J. Biol. Chem. 272:8090-8096(1997).

RT [2]

RN SEQUENCE FROM N.A.

RC TISSUE=heart;

RX MEDLINE=9704087; PubMed=883817;

RA Tanaka T., Inazawa J., Nakamura Y.; "Molecular cloning of a human cDNA encoding putative cysteine protease (PRSC1) and its chromosomal assignment to 14q32.1." Cytogenet. Cell Genet. 74:120-123(1996).

RN TISSUE=Brain;

RA Strausberg R.; Submitted (May-2001) to the EMBL/GenBank/DBJ databases.

RN [4]

RP SEQUENCE OF 18-24 AND 324-330, AND PROCESSING. Pubmed=108595;

RA Chen J.-M., Fortunato M., Barrett A.J.; "Activation of human prolegumain by cleavage at a C-terminal asparagine residue." Biochem. J. 352:327-334 (2000).

RN [5]

CHARACTERIZATION OF ACTIVITY ON ASPARTATE BONDS. Pubmed=8821970;

RA Halton S., Patel S., Vega F., Zurawski S., Zurawski G.; "Autocatalytic activation of human legumain at aspartic acid residues." FEBS Lett. 438:114-118 (1998).

CC -!- FUNCTION: Has a strict specificity for hydrolysis of asparaginyl bonds. Can also cleave aspartyl bonds slowly, especially under acidic conditions. May be involved in the processing of proteins for MHC class II antigen presentation in the lysosomal/endosomal system.

CC -!- CATALYTIC ACTIVITY: Hydrolysis of proteins and small molecule substrates at -Asn-| -Xaa-bonds.

CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).

CC -!- TISSUE SPECIFICITY: Ubiquitous. Particularly abundant in kidney, heart and placenta.

CC -!- PTM: Glycosylation.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires license agreement (See [http://www.isb-sib.ch/announce/NCBI_TaxID=6183/](http://www.isb-sib.ch/announce/NCBI_TaxID=6183;)).

CC -----

DR EMBL; YP09862; CAA0989_1; DR EMBL; D55696; BAA09530_1; DR EMBL; BC008004; AAH08004_1; DR MEROPS; C13_004; DR Genew; HGNC:9472; DR InterPro; IPR001096; Legumain_Pfam; PR00776; HEMOGLOBINASE. FT SIGNAL_1_17 CHAIN_18_323 LEGUMAIN_17 PROPEP_324_433 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD_91_91 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD_167_167 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD_263_263 N-LINKED (GLCNAC. .) (POTENTIAL). FT CARBOHYD_272_272 N-LINKED (GLCNAC. .) (POTENTIAL). FT ACT_SITE_148_148 POTENTIAL. FT ACT_SITE_189_189 CLEAVAGE (AUTOMATIC). FT SITE_323_324 N-D_Q_S LOSS OF AUTOACTIVATION. FT MUTAGEN_323_323 CONFLICT_31_31

SQ	SEQUENCE	433 AA:	49411 MW:	081AD2D0584E72A CRC64;
	Query Match	33.7%	Score 828;	DB 1;
	Best Local Similarity	38.2%	Pred. No. 4e-51;	
	Matches 179;	Conservative 80;	Mismatches 157;	Indels 52;
	Gaps 13;			
QY	8 FQILVFLHALLIFSAESRKQTQLNDVESSDKSAKGTWAVLVAGSNSEYYNRHOAIDC	67		
Db	3 WKVAYFELSVALTGIGA-----VPIDDPEGGKHVVWVSNWYRNROAIDC	50		
QY	68 HAYQILRKGGGLKDENTIVMYDDIAFSSSENPRCGVIIINKPDGEDVYKGPKDTYKEAVNV	127		
Db	51 HAYQIHLRNGIPIDEQIVMMYDDIAYSEDNPTCPIVINRPNGTDVQCVPKDTYGEDVTP	110		
QY	128 QNFYVNLLEGNESGVTG-SNGKVVKSGPNNDIYIYADHGAPGLJAMPITGEVNAKDFNVEW	186		
Db	111 QNFIAFLVRGDAEAVKGIGSKVLUKGPGDHFVLYFTDGSTGLVFP-NEDLHVKDNET	169		
QY	187 LEKMHKKRKYKNKAVTYVEACEGSGMFEGILKKLNNTYVATAANSKESSWWGVYCPESYPP	246		
Db	170 IHWYKHKMYRKAVFYIEACESGMMNH-LPDNINVATTAANPRESSVACYDE---	223		
QY	247 PSEIGTCIGDTFSISWLIEDSDHDMSKETLEQHQYHVVKRVRGSDVPESTHVCRGTEKML	306		
Db	224 --KRSTYLGDWYSVNWMEPSDVEDLTKETLHKQYHVKSHT---NTSHVMOYGNKTS	276		
QY	307 KDLSSYIGRNPDENDFTTESSS---PINSGLVNPNDPLPPLYQQRKTOKAPMGSL	361		
Db	277 TMKVMQFGQ-----MKRKASSPVLPVPLVHDLDLTPSPDVPLTMIRKLMTN--NDL	325		
QY	362 -ESKEAQAKKLDEKNHRKQDQSITDLRLSPSQTQNVLNLLSTRTGQPLVDDWDCFKT	420		
Db	326 EESRQLETEIQRHLDARHIEKSVRKIVSLLAASEAEVQLLSERA---PLTCH-SCPE	381		
QY	421 LYNESEKHNHC---GATVHYGLKYTGALANICNMNGVDYQTVSALEQAC	464		
Db	382 ALLHERTHCFNNHSPTYEXALRHLVLYNLCEKPYPLRHIKLSMDHVC	429		
	RESULT 12			
	HGLB_SCIMA			
ID	HGLB_SCIMA	STANDARD:	PRT;	429 AA.
AC	P09441;			
DT	01-MAR-1989 (Rel. 10, Created)			
DT	01-FEB-1995 (Rel. 31, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Hemoglobinase precursor (EC 3.4.22.34) (Antigen SM32).			
OS	Schistosoma mansoni (Blood fluke).			
OC	Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigida; Schistosomatidae; Schistosomatidae; Schistosomidae.			
OX	NCBI_TAXID=6183;			
[1]	SEQUENCE FROM N.A., AND SEQUENCE OF PRT; 429 AA.			
RN	RP SEQUENCE FROM N.A.			
RX	RX MEDLINE=89261925; PubMed=2725581;			
RX	RX MEDLINE=50342941; PubMed=2383765;			
RA	RA el Meanahy M.A., Ajit T., Phillips N.F.B., Davis R.E., Salata R.A., Malhotra I., McClain D., Akawa M., Davis A.H.			
RA	RA "Definition of the complete Schistosoma mansoni hemoglobinase mRNA sequence and gene expression in developing parasites," Am. J. Trop. Med. Hyg. 43:67-78 (1990).			
RT	RT [2]			
RT	RT SEQUENCE FROM N.A.			
RN	RN MEDLINE=89261925; PubMed=2725581;			
RX	RX Klinkert M.-O., Felleisen R., Link G., Ruppel A., Beck E., RT "Primary structures of Sm31/32 dihydroproteins of Schistosoma mansoni and their identification as proteases.";			
RA	RA Mol. Biochem. Parasitol. 33:113-122 (1989).			
RN	RN [3]			
RP	RP SEQUENCE OF 77-429 FROM N.A.			
RX	RX MEDLINE=87308326; PubMed=3305515;			
RA	RA Davis A.H., Nanduri J., Watson D.C., Cloning and gene expression of Schistosoma mansoni protease."			
RT	RT J. Biol. Chem. 262:12851-12857 (1987).			
RL	CC -!- FUNCTION: THIS PROTEASE IS USED BY THE PARASITE FOR DEGRADATION			

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C13.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC DR EMBL: Y07596; CA68871_1;

DR EMBL: AF022913; AAB81597_1;

DR EMBL: BC02037; AAH2037_1;

DR Genew; HGNC:8965; PICK;

MIM: 603087;

DR METOPS; C13_005;

DR InterPro: IPR001096; Legumain.

DR PFAM: PF01650; Peptidase_C13; 1.

DR PRINTS: PR00776; HEMOGLOBINBASE.

KW Hydrolase; Thiol protease; GPI-anchor biosynthesis.

FT ACT_SITE 164 POTENTIAL.

FT ACT_SITE 206 POTENTIAL.

FT CONFLICT 1 4 MAVT->SIHEA (IN REF. 1).

FT SEQUENCE 395 AA: 45251 MW: AF706DDAD13EFB2 CRC64;

SQ Best Local Similarity 23.4%; Score 297.5; DB 1; Length 395;
 Matches 95; Conservative 82; Mismatches 170; Indels 59; Gaps 14;

Qy 14 LHLILISAESRKTLQNLNDVESSDKSAKGTRAVLYAGSNEWYYRHQADICHAOIL 73

Db 14 LATVILISFGSVAASHTD-QAEQFRSGTINNAYVLTSRFVNRYREVENTLSVRSV 72

Qy 74 RKGGLDENITIVMDYDDIAFSENPRPGVLINKPDGE-DVY-KGVPKDYTKEAVNVONY 131

Db 73 KRGIPDSHVIMLAADMCRPNNPKPATVFSKRNMLNVYGDVEVDYRSYEVTVENFL 132

Qy 132 NVLGNESGVTGGNGKPNPNDNIFTYADHGAPGLIAMPCTDEVMAKDFNEVLENMH 191

Db 133 RVTGTGRPPST-PRSKRLLSDDRSNLILMTGHGNGELKFQDSEETNIELADAFOMW 191

Qy 192 KRMKYNKVVIVTEACESGSMPFGLRKRNLYATAANSKESSGVYCPESYPPPSETG 251

Db 192 OKRKNHSLFLIDTCGASMYERFSP-NIMAAASSOYGEDSU-----SHQDPDA-IG 242

Qy 252 TCGDTFSISLD-SDLHDMSKETLEQQYHVVKRRVGSVPETSHVCRFTGTEKMLKDYL 310

Db 243 VHLMDRTFYVLEFLBINPASQTNMNDLFQVCPSLCVSTP-----GHR 287

Qy 311 SSVTGRIPENDNFTFESFSPINSGLVNPRTDPLLYQORKICKAPGSLEEKACKKL 370

Db 288 TDLFQDPK--NVLTDFFGS-----VRKVEITITETIKLQQDSE 324

Qy 371 LDEKHNHKR-QIDQSTIDILRLSVKOTNVNLNTSTRTTQPLVDDW 415

Db 325 IMESSYKEDQMDKIMEPLIYA-EQLPVQAIIHQ----KPKLKDW 364

Search completed: May 27, 2003, 15:26:51

Job time : 17 secs